8-2010

Listening to students: The Lived Experience of Students Taking an Accountability Test

Laura Rutherford Crisp
lcrisp@utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Laura Rutherford Crisp entitled "Listening to students: The Lived Experience of Students Taking an Accountability Test." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Educational Psychology and Research.

Katherine Greenberg, Major Professor

We have read this dissertation and recommend its acceptance:

Howard Pollio, Gary Skolits, Ralph Brockett

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:
I am submitting herewith a dissertation written by Laura Rutherford Crisp entitled "Listening to Students: The Lived Experience of Students Taking an Accountability Test." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Educational Psychology and Research.

Katherine Greenberg, Major Professor

We have read this dissertation
And recommend its acceptance:

Howard Pollio

Gary Skolits

Ralph G. Brockett

Accepted for the Council:

Carolyn R. Hodges
Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

LISTENING TO STUDENTS:
THE LIVED EXPERIENCE OF STUDENTS TAKING AN ACCOUNTABILITY TEST

A Dissertation
Presented for the
Doctor of Philosophy Degree
The University of Tennessee, Knoxville

Laura Rutherford Crisp
December 2010

Dedication

This dissertation is dedicated to my son. His experience has opened my eyes to how important it is to listen to those who are most affected by educational reform: educators and students.

"..the educational function does not rest upon our ability to control, or our will to instruct, but upon our human nature and the nature of experience."
Dennison, 1969

Acknowledgements

My sincere thanks go out to the students at the recreation center who were willing to interview with me, as well as the head honcho, who is such a positive force at the center.

I also want to thank Kathy Greenberg who convinced me to get my Ph.D. by telling me that all I have to do is write a dissertation. Your supportive advice and kind but tenacious editing of my paper after such a long hiatus was much appreciated.

My gratitude extends to the rest of my valuable committee: Dr. Howard Pollio, Dr. Ralph Brockett, and Dr. Gary Skolits for all your advice, support and constructive feedback of my research.

I would also like to thank all of the research team members of the Applied Educational Psychology research team at the University of Tennessee for their help and insight in analyzing and interpreting the data for this project.

Last but not least, I would like to thank my family and friends who continued to believe I would finish my dissertation even when I myself did not. I especially want to thank my parents for all of their support during these last few months. I appreciate Ethan calling it my “desertation”, as it always made me smile. Finally, I want to thank everyone who continued to ask the annoying yet much needed question over these last three years, “When are you going to finish your dissertation?”

Abstract

The purpose of this study is to understand the lived experience of students taking a high stakes test. The phenomenological method developed by Howard Pollio (Pollio, Graves, and Arfken, 2005, Thomas and Pollio, 2002, Pollio, Henley, and Thompson, 1997) at the University of Tennessee was utilized to explore the perceptions of the experience of fourth and fifth grade students who took the Tennessee Comprehensive Assessment Program (TCAP) Achievement Test, an end of the year criterion-referenced, standardized achievement test given to students in Tennessee.

Nine students from two schools in East Tennessee were interviewed about their experience of taking the test. Based on the student's own words, a Ground and three themes were identified: (Ground) the test as different from other tests; *It just felt like another test, but like more important/ weird*; (Theme 1) reactions to the actual test: *It was kind of hard and kind of easy*; (Theme 2) strategies that help you on the test: *Yeah, you have to really try hard*; and (Theme 3): The purpose of the test: *It decides, well, I don't know, but it seems…..*

This study shows that asking children about their experience can enlighten our understanding of standardized testing practices that impact test scores schools rely on so heavily to prove their efficacy as an institution. If we are to support our teachers and our schools in this era of high stakes, we should acknowledge the voices of our students who are taking them, and adjust our policy and practices accordingly. Considerations for practice and recommendations are also included.

Table of Contents

List of Figures

Chapter I

Background

"Educators need to know what happens in the world of the children with whom they work" (Freire, 1998, p72).

A common declaration for education reform includes the need for improved student achievement. Whether it is to compete in a global market, to overcome racial disparity, or to become successful in today's information age, students are at the center of all educational reform. However, while students seem to be the focus, efforts are centered almost exclusively on the adult's role as reformers (Corbett & Wilson, 1995). Fullan (1991), a noted theorist on education reform, comments that students are regarded as only "beneficiaries" of reform and never thought of as "participants" (Fullan, p. 170).

In 2001, the No Child Left Behind (NCLB) act was completed in a reform effort to strengthen the accountability of K-12 education in the U.S. This landmark federal legislation requires states to develop assessments in basic skills in order to receive federal funding for schools. Signed into law in 2002, the bipartisan NCLB legislation requires annual testing for students in grades 3-8 and all students to be proficient in reading, writing, and math by 2014. Schools are required to report the scores of these annual tests to the public, disaggregating the data so that minority scores can be observed alongside the scores of the majority. If scores do not progress from year to year, schools are placed in categories and sanctions are given according to their label.

These sanctions can include state take-over, teacher and administrative firings, and school closure.

As of 2006, four years after the No Child Left Behind legislation was passed, many schools have been given the Needs Improvement label, a label given to schools who have failed to meet improvement scores for two years. A recent report calculating the number of these schools found that 22,873 schools did not make Annual Yearly Progress (AYP). This number represents 25.8 percent of all public schools or an increase of 1,699 schools from the 2004-05 school year (National Education Association, 2006). These numbers are significant suggesting that all districts, schools, and teachers have felt the impact of NCLB.

However, what do we know about the impact on students? Children are not often asked about their feelings, attitudes or thoughts about assessment. Instead, evaluation of children is often isolated to behavior changes or test scores. “Children have been denied both a voice and an essential feature of human identity, a rational standpoint” (Hendrick, 2000, p. 38).

It is not clear why we do not perceive students as having the capacity to be either actors in education reform or in having them inform its effects. The answer may be in the way we have conducted child research. As Kessen (1979) points out, “No other animal species has been cataloged by responsible scholars in so many wildly discrepant forms that a perceptive extraterrestrial could never see as reflecting the same beast” (p. 815). Since G. Stanley Hall began studying children using surveys, the scientific observation and classification of children has become the norm. Even today, at the University of Tennessee for example, the childcare on campus is contains the word “labs” in its name,

with one-way glass walls and microphoned rooms in order to observe children without them knowing. This is just one example of many that demonstrates the mainstream view of children in education; scientifically measurable, separate from adults, and easily manipulated to change (Smith, 1983).

If you look at the word accountability, it is defined as, “making teachers and schools responsible for student learning, usually by monitoring learning with high-stakes tests” (Woolfolk, p. 540, 2007). Test-based accountability reform remains hotly debated among educators and researchers but has found widespread support among policy-makers and the business sector (American Psychological Association, 2001, Amrein and Berliner, 2002, Darling-Hammond, 1991, Figlio, 2006, Kohn, 2000, Madaus and Clark, 2001, Popham, 2001). There have been a number of accountability reform movements in the United States with NCLB being the first to require such high stakes, or consequences associated with test results. The emphasis on high stakes accountability is the result of growing public distrust that schools are providing a good education. Accountability movements have also gained momentum by the growing acceptance that U.S. superiority in technological advancements is losing its foothold.

One other belief has encouraged the national feeling that accountability is a necessity: the belief that our economic welfare is dependent upon the quality of our educational system. In *The Guide to Education and No Child Left Behind*, it states that “raising student achievement directly leads to national economic growth” (US Department of Education, 2004, p.I) . While no statistics were offered in qualifying this statement, it is certainly a belief that has been echoed by many politicians and business people who not

just promote, but see that the accountability movement is essential for U.S. growth and economic stability.

However, businessman and politicians are not the only ones who favor accountability reform. Groups such as the National Council of La Raza, the Citizens' Commission on Civil Rights, and the National Center for Learning Disabilities support No Child Left Behind because it promotes equity in education as well as improving transparency. The public is also concerned about education and high stakes testing enjoys a high approval rating (Shirley, 2009).

The accountability movement has also run parallel to the higher standards movement. The higher standards movement is a response to a growing understanding in education that increasing expectations of children will prepare students for college and the workforce (Duncan, 2009). The higher standards movement has altered many of the accountability tests that are given to students each year. For example, next year, Tennessee will be altering their tests to align with higher standards. As it states on the Knox County Schools website, "The Knox County School System is redefining standards and raising the bar in four core areas to better prepare students to compete in a shrinking and increasingly competitive world" (Knox County Schools, 2009).

The higher standards movement proposes that currently, standardized tests are only testing basic, low level knowledge and standards. The standards– as well as the tests - should be higher. The higher standards movement also seeks to improve efforts to synch the curriculum with the tests (Duncan, 2009).

Massachusetts, a state that is leading the higher standards movement, has a more rigorous end of the year exam that is aligned to the curriculum. On the surface, the

premise behind the accountability movement as well as the standards movement does not seem flawed or out of synch with what we know about teaching and learning. High expectations should yield higher attainment of achievement (Madaus, 1991).

Statement of Problem

Education reform that uses accountability to drive change makes several assumptions about the effects of testing 1) testing will improve the educational system, 2) instruction will improve 3) tests will measure necessary outcomes of the school, and 4) tests will measure school/class/student success (Madaus, 1991). However, since the NCLB law has been passed, some educational researchers, teachers, and administrators have been critical of a test-driven accountability system (Amrein and Berliner, 2002, Darling-Hammond, 1991, Figlio, 2006, Kohn, 2000, Madaus and Clark, 2001, Popham, 2001). The expanding use of accountability tests and the lack of discussion between policy makers and those who implement the policy have certainly fueled this backlash against standardized testing. Additionally, the conceptual assumption that accountability testing makes - that one can reduce student achievement to a number - is problematic for many who reflect on the complex issue of what constitutes real learning (Kohn, 2000).

Critics of testing and its effects repeatedly cite changes occurring from accountability measures like NCLB that have a direct or indirect impact on children. Popham, a long time educator and test developer believes that high stakes tests are doing serious harm to children. In *The Truth About Testing: An Educator's Call to Action*, Popham (2001) details how current testing policies result in negative consequences in the classroom, including a focus on test scores instead of actual learning,

misidentification of effective and ineffective schools, and reduction of curriculum to cover only test content (2001). Popham (2007) is also concerned that all but a few of the tests tied to high stakes are instructionally insensitive, meaning the test scores from these tests do not reflect how well the students are taught. The variance in student scores on these tests have more to do with socio-economic factors than quality of instruction (Popham, 2007, Nichols & Berliner, 2005).

It has been repeatedly recognized among researchers that an accountability system *solely* based on a high stakes test is problematic (Kogan, 1986, Hirsch, 1996, Figlio, 2006, Darling-Hammond, 1989, Brennan, 2001, AERA, 2000, etc.). Linda Darling-Hammond (2009) labels our current culture of accountability as a 'Bureaucratic Approach' to education. "The fundamental assumption is that this process, if efficiently administered, will produce the desire outcomes" (p. 47). Efficient administration is played out in an ever increasing role of administrative controls in the classroom that affects curriculums, classroom schedules, and testing regimens (Darling-Hammond, 2009). The more bureaucratic a system, the less control teachers have to alter instruction to meet the needs of students. Teaching, simply put, cannot be standardized, and when it is, it stifles learning and innovation in teaching. "Not only is bureaucratic management at odds with innovation, it is substantially at odds with student learning" (Darling-Hammond, 2009).

There is promising evidence of decentralization improving learning outcomes and encouraging innovation. Finland decentralized their educational system due to high unemployment caused by a shift in their economy. The goal was to allow a space for innovation in education that could develop students to meet the needs of the 21st century.

Student outcomes have risen and Finland ranks in 1st and 2nd place in literacy, math, and science (Hargreaves, 2009).

The accountability movement puts pressure on teachers and administration, which can be a good thing. However, researchers who study educational reform believe it is an unexceptional way to reform education because it does not allow leadership to actually lead nor does it provide administrators or teachers sufficient support (Spilane, 2009, Levin, 2009). It is also not a sufficient strategy to produce better learning outcomes (Levin, 2009). Levin (2009) purports that for reform to be effective it must include the following: "changing teaching and learning (and assessment) practices, improving teachers' skills to do so, strengthening leadership capacity, improving student engagement in learning, and reaching out to parents to support their children's learning (p. 262). High stakes testing does not encourage or allow for the type of whole-system improvement required for real change, nor does it provide a positive environment or motivation required for inspiring change (Fullan, 2009).

Another criticism of high stakes testing is its negative impact on teaching and learning (Madaus, 1988; Nolen, Haladyna, and Haas, 1992; Birkmire, 1993; Darling-Hammond, 1991). Test preparation for the standardized tests begins in the fall and heavy test preparation six-weeks prior to the test in the spring is commonplace. Schools tend to teach to the test by reserving the majority of instruction time to be devoted to reading and math. Less time is reserved for other subjects such as social studies and science, resulting in a narrowing of curriculum (Amrein and Berliner, 2002, McCracken and McCracken, 2001). Preparing for the tests involve a higher number of pre-tests and benchmark tests that take instructional time and resources to administer and are not

focused on learning (Jones et al., 1999, Barksdale-Ladd and Thomas, 2000, Darling-Hammond and Wise, 1985).

NCLB's intention is to increase student achievement by increasing the level of accountability using high stakes tests as the measure of both student learning and teacher efficacy. However, evidence suggests NCLB is both successful and unsuccessful. While a large percentage of schools are reporting increases in achievement as shown by their yearly achievement scores, this progress does not seem to transfer when compared to the National Assessment of Educational Progress (NAEP). The NAEP is the national norm-referenced assessment given to a percentage of students across the nation. NAEP is a test used as the nation's benchmark when keeping tabs on student achievement. For example, increases in the Texas Assessment of Academic Skills (TAAS) did not transfer to increases in the NAEP (Klein, Hamilton, McCaffrey, and Stecher, 2000). When achievement scores improve based on NAEP data, they do not increase in reading scores (Amrein-Beardsley & Berliner, 2003).

Another study, analyzing 28 states with high stakes found there was no consistent evidence that academic achievement increased (Amrein and Berliner, 2002). Despite the federal government's insistence on using standardized tests as a proof of teacher efficacy, teachers and administrators disagree that the tests can provide the desired results. "Over and over again, teachers and administrators discounted the validity of high stakes test scores as an accurate reflection of the teaching and learning which are occurring in schools" (Daniels, 2002, p.83).

The negative impact that accountability testing has on poor and minority students is also a concern among researchers. The major tenet in NCLB is to close the gap

between minority and majority achievement, therefore heightened accountability was intended to increase minority achievement. While standards have equalized for minority students, Darling-Hammond (2007) argues that access to the resources needed for educational success have not materialized. Amrein and Berliner (2002) explain that minority students and students with low socio-economic status are affected by the repercussions of high stakes testing more so than white students. For example, a Tennessee inner-city school with higher percentages of minority and low socio-economic students designs their curriculum around the end of the year test, according to the principle of a school in the district where this study took place, narrowing the curriculum more than other schools that typically do well on the test (Elisa Luna, Personal Communication, May 8, 2008). At the high school level, there is strong evidence to suggest that a high stakes testing environment encourages kids who perform at lower levels to drop out (Heilig & Darling-Hammond, 2008).

Researchers have also been concerned with testing's impact on student motivation. There is evidence for and against the claim that high stakes tests are motivating. Some studies have shown an increase in effort and/or retention of information under a high stakes testing scenario, specifically among low achieving students (Roderick & Engel, 2001, Stecher, 2002). However, other researchers reported that high stakes testing decreased student motivation and found an increase in dropout rates among high school students (Amrein and Berliner, 2003, Madaus & Clark, 2001).

Few studies concerning the effects on students have considered the lived experience of students and what the tests mean to them. The importance of studying the impacts of testing on students is made well by Madaus and Clarke (2001):

> In many other areas where technology and policy intersect, the public insists on oversight-including technical oversight – to protect individuals from unintended negative effects. For example, faced with the policy decision to introduce a major new untried medical technology to millions of children, particularly a treatment that would be given to healthy children as well to those who were ill, the public would ask about the safety, efficacy, quality, and social and economic effects of the new technology or treatment, and public agencies have been established to address such concerns systematically. The effects of testing are now so diverse, widespread, and serious that it is necessary to establish mechanisms for catalyzing inquiry about, and systematic independent scrutiny of them. (p.22)

In studies done to investigate teacher's views of high stakes testing, teachers expressed concern for the effects the test had on their students (Adams, 2006, Crisp, Williams and Greenberg, 2006, McCracken & McCracken, 2001). Despite the amount of research reporting adult views on the effects of high stakes testing of children, there have been few studies that specifically address how students themselves experience high stakes tests. Of those studies, few have been done on elementary school children and none have been done in Tennessee. Therefore, understanding students' experience taking a high stakes accountability test in Tennessee is important when exploring the effects of accountability testing.

It has been found in recent research, that asking students about their school experiences can be used to strengthen reform efforts (Noguera, 2007, Corbett and Wilson, 2001). Corbett and Wilson (2001) found students to be useful "windows" through

which to view effects of reform. Not only were there responses “profoundly penetrating” they were able to describe not only effects of reform but their insightful analysis of it.

> My teacher know how to talk to you, like when you having a problem. Instead of having a temper or nuttin’, they just be nice. You can go to them and ask a question. They just don’t want to hurry you up and get you out of the class. My other teacher is always saying: Didn’t you hear me!? Didn’t you hear me!? I’m not repeating it! (p.1)

Ann Ferguson, author of *Bad Boys: Public Schools in the Making of Black Masculinity* discovered that talking to boys added significantly to her knowledge of the state of education and its effects on African-Americans. In a mixed methods approach of ethnography and qualitative interviews, Ferguson observed both the actions of boys and the teachers’ reactions to them. The significance she placed on these events as an adult was very different than the child’s view in many cases. “I assumed at the start that I would learn *about* kids; but it was not long before I was obliged to question this premise and begin to learn *from* children” (p.11).

For example, Ferguson was interested in the understanding of boys in the Punishing Room, the room where students were sent if their behavior warranted disciplinary action. Each student who went to the room had a file in which ‘incident reports’ were added each time. One day, a boy happened to see the thickness of his file. His reaction was different than what the researcher expected. While Ferguson expected a reaction of dismay, the boy was actually proud of the thickness of the file and wondered if he had the largest one. Thus, former research that focuses on the student’s experience

provides important discoveries in motivation and learning, discoveries that can inform research and practice in education.

Understanding children's experience is more than a national concern and has garnered greater importance in the area of children's rights. In 1989, the Children's Act was passed that made it a requirement to consult children when assessing a child's educational, physical or emotional needs (Greig & Taylor, 1999). After a United Nations convention that focused on the rights of the child, a proclamation signed by 61 countries states that young children's views and feelings should be taken into account when developing policies and providing services for them (MacNaughton, Smith & Davis, 2007).

Christensen and James (2000) feel there is a great need to listen to children "as social actors in their own right in contexts where, traditionally, they have been denied those rights of participation and their voices have remained unheard" (p.2). Sociologists and historians have embraced this concept in recent decades by including the experiences of children in their research. Listening to children is central to recognizing and respecting their worth as human beings. It is important to understand the view of children and how they perceive culture; otherwise, children are denied their standpoint as social actors in the making of history (Roberts, 2000). In the case of high stakes testing, much more could be lost if we don't ask children their experience. While education reform focuses on raising achievement, it may contain major implications for motivation and learning.

Listening to the voices of children has been identified as "the most neglected aspect of child development research" (Greig & Taylor, 1999, p. 81). Myth surrounding the subject of asking children about their experiences typically is focused on a child's

ability. These issues concern child's cognitive abilities, the validity of the statements and the interpretations of the statements by the researchers.

The phenomenologist Merleau-Ponty has argued that children's views of the world are unique. A child's consciousness has its own arrangement, one wholly different than the consciousness of an adult. Therefore, it cannot be compared to an adult's as somehow deficient or lacking, but a completely different phenomenon altogether (Rojcewicz, 1987). For example, a child's understanding of his body is different from an adults'. An adult sees one's body as something separate from their thoughts, like an object. A child, however, understands his body to be something that allows him to make contact with the world. As Rojcewicz (1987) writes:

> The child does not live in the scientific world, nor does he live in the world of ideas. He is neither a materialist nor an idealist. These are adult categories, and it is a mistake to translate the child's consciousness into them. The child's realm is instead that of the phenomenal, the lived, and it is that realm that he is expressing. (p.205)

Social researchers who make a case for researching children have looked at feminist work to draw parallels to the rationales they have developed on the importance of studying women (Roberts, 2000). For example, Graham's (1984) work suggests that traditional research methods like survey research have a tendency to 'fracture' the voice of women in social science. She suggests that a narrative tradition can remedy this problem (Graham, 1984).For example, in a study reviewed for this dissertation that surveyed elementary school children, fourth graders responded that testing improved critical thinking skills, and 8 out of 10 students believed that testing helped them apply

information instead of just recalling it. It is doubtful that a fourth grader would have an adult educator's understanding of terms like "critical thinking" or "apply", and "recall". (Debard, 2000). If the Debard (2000) study had followed up with interviews, the researchers may have had a better understanding of their answers.

An increasing need for research that allows children to communicate the effects of testing is necessary for us to gain a full picture of this policy. This dissertation attempts to explore the lived experience of students taking the high stakes Tennessee Comprehensive Assessment Program (TCAP), Achievement Test, taken at the end of the third and fourth grade years. Informal, phenomenological interviews were done at a free after-school program at a recreational center in Knoxville. We can fully understand the ramifications of accountability reform- not by seeking answers from school administrators, teachers, policy centers, or well-regarded think tanks - but by asking the students themselves.

Purpose of the Study

At this point, there have been few studies focused on exploring the experience of children taking a high stakes test and none in Tennessee. Prior research has shown that we can gain important understandings if we listen to children (Ferguson, 2001) and that It is important to engage children during educational reform (Corbett and Wilson, 2001). The purpose of this study is to explore the perceptions of the experience of fourth and fifth grade students who took the Tennessee Comprehensive Assessment Program (TCAP) Achievement Test, a state-mandated high stakes test given to all students in grades 3-8. The primary goal of the phenomenological method is to make explicit what is seen using

the participant's own words. Interviews that use an open-ended question as well as follow-up questions that encourage an open dialogue between researcher and participant is used.

Research Question

This research study is interested in exploring the experience of students taking a high stakes test. The question that was asked to students was "What was it like for you when you took the TCAP test?"

Delimitations

The main delimitations will be summarized in the following statements:

1. Since the purpose of phenomenological research is to determine a clear picture of the essence of the human experience related to some phenomenon, this study does not seek to explore all possible effects of testing on students.
2. Because of the methodological approach used, the study used a small sample size.
3. One site was selected within the city limits of Knoxville, Tennessee at a recreation center where students from two different schools spent their afternoons. One school, despite being an 'inner city school' won an award for their improvement in their TCAP Achievement Test scores, the other was in the needs improvement category at the time of the study and had just been labeled as a Title 1 school. The students were not selected based on pre-

determined race, sex, class classifications, nor did the study view the TCAP Achievement Test scores to determine whether or not the specific students were successful or unsuccessful at taking the TCAP Achievement Test. Thus, the study does not try to focus on a certain category of student however age was limited to 9-11 yrs.

Limitations

The main limitations will be summarized in the following statements:

1. The sample size will be limited to fourth and fifth grade students from two schools in Knoxville, TN. The study may not be able to take in account different aspects of children's experiences taking an end of year, high stakes test in larger or smaller locations within a different culture or social context.
2. The study will not focus on the generalizability of the findings but will look at statements that all students make in order to come to an understanding of the shared perceptions of these students of the experience of a taking a high stakes test.

Significance of study

The intent of this study is to provide information about what the experience is like for students who have taken the TCAP Achievement Test, the standardized, criterion-referenced, end of the year test. Results of this test are tied to high stakes that could lead to the reorganization of the school. As stated earlier high stakes testing reform makes

several assumptions: testing will improve the education system, instruction will improve, and tests measure necessary outcomes of school, and tests can measure school success. I anticipate that this study could contribute to the knowledge of elementary teachers and administrators in preparing for and giving children high stakes tests. This study could have an impact on how teachers respond to children before the preparation of the tests or during the actual testing week. Examining the nature of a learner's experience may allow other researchers, teachers, and school administrators to become aware of how their own assumptions and approaches affect students taking a high stakes test. In relation to research in teaching and learning issues surrounding high stakes tests, this study may provide insight into the experience of students beyond the two specific schools and what that experience may be like for other students of the same grade level in Tennessee.

Organization of the dissertation

Chapter I presented the introduction; statement of the problem; purpose of the study; and significance of the study. Chapter II contains the review of literature and research related to the effects of students taking high stakes tests, including the evolution of accountability; teacher's perceptions of children's response to testing; and literature on children's response to high stakes tests. The methodology and procedures used to gather and analyze data are presented in Chapter III. The results and analyses and findings from the study will be contained in Chapter IV. Chapter V will include a

summary of the study and a discussion of conclusions and recommendations for further research.

Chapter 2

Introduction

High stakes testing has been the driving force of the accountability reform movements over the past three decades. In this type of reform, tests are used to measure the quality of teaching and the level of student achievement. High stakes testing is a general term describing tests that are tied to consequences for teachers or students. High stakes can be tied to students, teachers, principles, schools, and districts. In the last eight years, high stakes testing, has been a federal government mandate; states must comply to continue to receive federal monies for education. The intention of No Child Left Behind is to improve student scores on end of the year tests, a term called Annual Yearly Progress (AYP). The goal of the legislation is to bring all student groups to a proficiency level in reading and math by 2014. Sanctions are made if progress is not made each year.

While accountability testing is now being used nationally, debates still continue on this issue. The primary question critics continue to ask is whether high stakes testing is improving teaching and student learning. Many researchers, activists, educators, and parents have asked policymakers to take a closer look at the use of high stakes testing and whether or not accountability using high stakes tests is really the tool to meet this goal (Madaus & Clarke, 2001, Manna, 2006, Gunzenhauser & Hyde, 2007).

In order to understand how accountability reform began and issues arising from the birth of federal accountability testing, it is necessary to understand its history. In

addition, an understanding of the types of tests that are being used to measure student progress can further build on or inform the socio-historical context of testing that influences children's experience. Thus, the literature review includes the following categories: the evolution of accountability, an overview of NCLB and the Tennessee Comprehensive Assessment Program (TCAP) Achievement Test, teachers' views of children's response to testing, and a review of studies done on children's response to high stakes testing.

A search was conducted initially in ERIC. Three searches were conducted for the three different subjects. The first search I used the thesaurus words: *history* and *high stakes tests*, as well as *educational history* and *high stakes tests*. I conducted a search on teachers using the search words *teacher attitudes* and *high stake tests*. Finally, I conducted a search for student studies using the words, *student attitudes* and *high stakes tests*. I also used the search words *achievement tests* in place of high stakes tests and the words *minority groups* and *minority group children* in place of student attitudes. I also conducted similar searchers on PSYCHInfo and Education Full Text.

However, many of the studies I found in the reference sections of the few articles that showed up in my initial database search. I was able to find these articles in the John C. Hodges library at the University of Tennessee, Knoxville or, if they weren't accessible in the library, I ordered them through ILLIAD.

The History of Accountability in Schools

The federal government has relied upon accountability testing to measure the efficacy of the programs it funds since the 1960's. However, high stakes testing has not

historically been a favorable way to assess schools (Mazzeo, 2001), nor an effective way to make programs accountable (Popham, 2008). It is important to understand the historical development of accountability, including political and social shifts affecting the accountability movement. The following is a summary of the major events of this movement.

Accountability testing has been a factor in measuring the efficacy of federal education programs since Johnson's Elementary and Secondary Education Act (Mantel, 2005). However, NCLB, a re-authorization of ESEA, put accountability testing as the main provision (Jennings, 2003). A number of historical factors influenced the testing movement: 1) federal governments efforts toward equality in education; 2) economy changes that questioned our nations international status; 3) reports that incited a perception of failing schools; 4) the reporting of a decline in NAEP test scores and scores comparing us unfavorably to our international rivals; and 5) changing employment demands (Macpherson, 1996, Jones, Jones, and Hargrove, 2003).

Public education grew significantly in the 50's, and its main challenge at this time was keeping up with demand, not quality. Schools enjoyed the status of being symbols of growth and catalysts for growing communities (Jones, Jones, and Hargrove, 2003). In 1958, however, national attention was focused on schools in the aftermath of the Russians' Sputnik launch, which led some Americans to believe that the U.S. might be slipping in dominance. In order to maintain this dominance, federal government focused on the local and state run public education system.

A national conference was held to brainstorm about better ways to prepare our children in science and math. Congress passed the National Defense Education Act

(NDEA) in response to the perceived need for better science education (Mantel, 2005). Title V of NDEA provided money to states for “guidance, counseling, testing, and identification of able students” (Spring, 2008, p.402). Also, some funding was provided for states to develop data-gathering and reporting systems (Spring, 2008).

In the 1960’s, Lyndon B. Johnson signed the Elementary and Secondary Education Act (ESEA) providing funds to school districts to help students in depressed rural and urban areas. ESEA was the first federal legislation to markedly change the federal government’s role in U.S. public education and student assessment (Anderson, 2007). ESEA’s intent was to bring about equality in education. One of ESEA’s most important provisions was Title I, which provided funding for programs focused on economically disadvantaged school children (Popham, 2008).

To ensure that federal monies were to be used for their intended purpose, Robert F. Kennedy fought to make certain the bill included evaluations to measure student achievement (New York State Education Department, 2006). Programs given funds were required to evaluate progress. However, because there was no consistent way in which people were collecting data and no consequences if a program was found making little progress, the evaluation of these programs was deemed ineffective (Popham, 2008).

While efforts in accountability assessment were being attempted because of Title 1 programs, efforts to implement statewide testing systems for school accountability largely failed (Linn, 2005). In the 1960’s, there was an actual debate among politicians on the effects of accountability. For example, the Citizen’s Action Committee, a group that proposed an accountability bill in California, wanted a testing system in place in order to set minimum standards of achievement and to be able to evaluate the California school

system. Dissenters of the bill argued that a state wide testing system would be a barrier to quality school programming, inhibit creativity, and would limit the curriculum in that it would force schools to homogenize their curriculums to meet state standards. Because testing relies on objective data, curriculum would then have to be minimized in order to conform to these tests. Instead, the few state-wide testing systems that were put in place were used for placement and student guidance (Mazzeo, 2001).

Another important event in the 1960's came from the Coleman Report, titled, *Equality of Educational Opportunity* published in 1966. It is important to note the historical context of Coleman's report. Coleman was commissioned by the government to investigate a conundrum developed by the reaction to the Title I grant money. Some believed that ESEA was making it economically advantageous for districts to keep the poor and minorities where they were in order to remain eligible for Title I grant money, retarding the push for integration. Coleman was to attempt to answer the question: Would compensatory education (ESEA) or integration improve student achievement within the disadvantaged populations? His conclusion was that neither promised positive results (NYS Education Department, 2006).

The Coleman report's importance, in terms of the development of accountability testing, was the idea of outputs. Coleman concluded that what students actually learned (output) was more important than inputs. Inputs included such things as teacher salaries, school building conditions, quality of equipment, and other resources (Coleman, 1972). High stakes testing's primary focus is on outputs.

Finally, in 1969, in response to reports of Title 1 fund misuse and critiques from scholars accusing Title 1 programs of not recording data that could prove the programs'

effectiveness, the federal commissioner at the time, James Allen, proposed a solution. The National Institute of Education (NIE) was opened to not only analyze the federally funded education programs in question, but to also study whether federal funds did in fact improve student performance in inner-city schools. Accountability, therefore, would be measured as academic achievement (NYS Education Department, 2006). Allen also started the National Assessment of Educational Progress (NAEP), a national system of tests designed to track student achievement throughout the country over time. His intention was to help states see where they might need to improve (Federal Educational Policy, 2006).

The 70's brought economic challenges and the worry that government could not keep up with the costs of social programs (Macpherson, 1996). With this financial downturn, criticisms of schools and their ability to educate took root. Lack of quality education was believed to be a factor in the decline of the financial climate (Macpherson, 1996). Studies criticizing student achievement produced an even greater need for reform in education (Mantel, 2005). Ironically, the conservative administrations of Nixon, Ford, and Reagan all asked for cuts in education as none of these presidents were in favor of an increase in the federal role of education. However, Title 1 funding (a.k.a. ESEA) maintained its bipartisan support because of its favorable distribution. The reason behind this support was that virtually all districts received the Title 1 money (DeBray, 2006).

In 1974, concerns about the lack of accountability of Title I funds spurned another reauthorization of ESEA. The Title 1 Evaluation and Reporting System (TIERS) was created during this time and was a success in that it was able to provide data to congress on the Title 1 programming and it created the infrastructure in which local and state

education agencies could be trained in assessment and reporting practices in order to improve services (Reisner et. al, 1982). TIERS contributed to the growing use of norm-referenced tests in this era (Linn, 2000). Norm-referenced tests are tests that compare a student's score against the scores of a group of individuals who have already taken the exam, also called the norm group. Norm-referenced tests are designed to highlight achievement differences between and among students to produce a dependable rank order of students across a continuum of achievement from high achievers to low achievers (Stiggins, 1994).

The use of these norm-reference tests however drew criticism from some researchers on several different aspects of the program. Some researchers were concerned that the norm group used did not reflect the socio-economic and racial diversity of the test-takers (Linn, 2000). Since mostly poor and minority populations were being tested, critics felt that the individuals taking these tests should be compared to a representative norm. Also, claims were made that test items were not aligned with learning objectives of Title 1 students. Instead of measuring effects of a specific program, the tests measured general abilities as is indicative of norm-referenced tests (Linn, 2000).

Inconsistency of the types of tests used across states continued to be a problem when making state-to-state comparisons. When different tests were used for different programs, comparisons couldn't be made (Popham, 2008). Also, critics cited inconsistent administration of tests as a major inhibitor of the validity of the test scores. Standardized tests, as their name describes, are meant to be given in a predetermined, standard way to ensure that the students taking the test are taking it under the same conditions as the norm group (Popham, 2008).

The impetus behind school change was one of the most understated but momentous shifts in education reform in the late 1970's. During this time, business leaders became the motivating force behind school change, replacing local community activists as primary change agents in the previous decade. The beginning of federally funded business-school partnerships/grants prompted this alteration (Federal Education Policy, 2006).

The growing perception of a declining public educational system was pushed further into the limelight in the controversial report *A Nation at Risk: The Imperative for Education Reform* (Macpherson, 1996). Published in 1983 by the U.S. Department of Education's National Commission on Excellence in Education, the report used test scores to show the mediocrity of public education. The list below outlines the evidence that the commission used to make its argument:

- International comparisons of student achievement, completed a decade ago, reveal that on 19 academic tests American students were never first or second and, in comparison with other industrialized nations, were last seven times.
- Some 23 million American adults are functionally illiterate by the simplest tests of everyday reading, writing, and comprehension.
- About 13 percent of all 17-year-olds in the United States can be considered functionally illiterate. Functional illiteracy among minority youth may run as high as 40 percent.
- Average achievement of high school students on most standardized tests is now lower than 26 years ago when Sputnik was launched.

- Over half the population of gifted students do not match their tested ability with comparable achievement in school.
- The College Board's Scholastic Aptitude Tests (SAT) demonstrate a virtually unbroken decline from 1963 to 1980.
- Average verbal scores fell over 50 points and average mathematics scores dropped nearly 40 points.
- College Board achievement tests also reveal consistent declines in recent years in such subjects as physics and English.
- Both the number and proportion of students demonstrating superior achievement on the SATs (i.e., those with scores of 650 or higher) have also dramatically declined.
- Many 17-year-olds do not possess the "higher order" intellectual skills we should expect of them. Nearly 40 percent cannot draw inferences from written material; only one-fifth can write a persuasive essay; and only one-third can solve a mathematics problem requiring several steps.
- There was a steady decline in science achievement scores of U.S. 17-year-olds as measured by national assessments of science in 1969, 1973, and 1977.
- Between 1975 and 1980, remedial mathematics courses in public 4-year colleges increased by 72 percent and now constitute one-quarter of all mathematics courses taught in those institutions.

- Average tested achievement of students graduating from college is also lower. Business and military leaders complain that they are required to spend millions of dollars on costly remedial education and training programs in such basic skills as reading, writing, spelling, and computation. The Department of the Navy, for example, reported to the Commission that one-quarter of its recent recruits cannot read at the ninth grade level, the minimum needed simply to understand written safety instructions. Without remedial work they cannot even begin, much less complete, the sophisticated training essential in much of the modern military.

  (Gardner, et.al., 1983, pp. 8-10)

Shepard (1993) notes that after the publication of *A Nation at Risk*, a shift in the educational reform movement occurred, in that the relationship between our educational system and the economic strength and competitiveness of our nation was inexorably intertwined.

In 1988, ESEA was reauthorized once again and was focused on altering accountability requirements. The bill required states to define target achievement requirements and to assess students on a yearly basis (many states were testing biannually or even less infrequently). The reauthorization also promoted states to use criterion-referenced tests and move away from norm-referenced tests (Popham, 2008). These two tests are different in the way the grades are analyzed and also in what is being tested.

Popham (1981) defines a criterion-referenced purpose: “A criterion-referenced test is used to ascertain an individual’s status with respect to a defined behavioral domain” (p.27). This defined behavioral domain specifies in detail what behaviors or knowledge are required in order to be deemed proficient. Criterion-reference tests are then graded on the basis of whether a student has reached a satisfactory or unsatisfactory level. For example, in the TCAP Achievement Tests, students are measured against a Performance Index which has three levels; Advanced, Proficient, Below Proficient.

On the other hand, a norm-referenced test measures a student based on other student’s performance. Scores of one student is compared to what is called a norm group, or a group of students that have previously taken the test. The scores are given as percentiles. For example, a student who gets a 91% on a reading comprehension norm-referenced test means that the student scored as well or better than those in the norm group.

Another difference of norm-referenced test is in what the test measures. Norm-referenced tests more generally define the area in which it is measuring. However, a criterion-reference test would clearly define the specific aspects of reading comprehension it is measuring. If five aspects of reading comprehension were defined, then the test would have five groups of questions in order to measure the performance of an individual in each aspect. The data taken from the test would then theoretically be able to show which aspects of reading comprehension an individual had mastered (Popham, 1981). The growing use of criterion-referenced tests was seen as a positive trend in testing and measurement that could better measure individual student achievement.

The next growth spurt for accountability testing came in the form of a list of goals. Goals 2000 was a list that began to develop in 1989, at the nation's first education summit attended by the nation's governors and former president George H.W. Bush (Austin, 2000). Out of the summit was a promise to develop educational goals that were made public in 1990 (Finn, 1990). The goals were originally coined America 2000 but its name changed to Goals 2000 during the Clinton administration and are listed below:

By the year 2000:

1. Every child will start school ready to learn.
2. The high school graduation rate will increase to at least 90 percent.
3. American students will leave grades 4, 8, and 12 having demonstrated competency over challenging subject matter including English, mathematics, science, foreign languages, civics and government, economics, art, history, and geography; and every school in America will ensure that all students learn to use their minds well, so they may be prepared for responsible citizenship, further learning, and productive employment in our nation's modern economy.
4. The nation's teaching force will have access to programs for the continued improvement of their professional skills needed to instruct and prepare all American students for the next century.
5. U.S. students will be first in the world in science and mathematics achievement.
6. Every adult American will be literate and will possess the knowledge and skills necessary to compete in a global economy and exercise rights and responsibilities of citizenship.

7. Every school in the United States will be free of drugs, violence, and the unauthorized presence of firearms and alcohol and will offer a disciplined environment conducive to learning.
8. Every school will promote partnerships that will increase parental involvement and participation in promoting the social, emotional, and academic growth of children.

   (Austin, 2000)

These goals were not just developed by government but heavily influenced by business interests. Just before the summit, U.S. students had just been ranked 14th in math and 13th in science among other industrialized countries (Bierbauer, 1996). Raising the standards for American education was a perceived need to raise our standing in educational achievement and thus secure our place in the economic global market (Hansen, 2006). Out of the summit, the idea that raising the standards used to measure achievement was the answer. This would give businesses the job pool they needed to meet the higher demands of the positions they promised to create.

Additionally, the summit was not an open affair. Members of the business community and mostly conservative politicians were invited. Excluded at this meeting were the important voices and opinions of teacher and parent groups. These groups were not invited to the meeting (Bierbauer, 1996).

Clinton, a proponent of standards and assessment, used the Goals 2000 to launch a national campaign enticing states to adopt a standards-based curriculum with a far reaching influence: 48 states developed standards and assessments based on this initiative (Hansen, 2006). Thus, Goals 2000 allowed the federal government to play a

much larger and more influential role in education using assessments as their tool for change.

However influential the last ESEA reauthorization in 1988 and Goals 2000 was, Bush's reauthorization of ESEA, the No Child Left Behind Act (NCLB), marks a historical turning point in the federal role in accountability reform (Jennings, 2003).

No Child Left Behind

No Child Left Behind passed in 2002 markedly increased federal power over education by requiring states to test *all* grades in each grade range, 3-5, 6-9, and once during high school. NCLB tied stiff penalties to states if they did not comply with the new provisions of the Act. Scores were also required to be disaggregated by race and more serious consequences were tied to evaluation of these scores. NCLB also required state participation in the National Assessment of Educational Practice (NAEP) in 4^{th}, $8^{th,}$ and 12^{th} grades in math and reading while prior to NCLB, participation was voluntary.

All states were required to set their own provisions on the consequences or rewards for these individuals who are in charge of a classroom, school, or district that is not meeting Annual Yearly Progress (AYP) on the standardized achievement tests (ASCD, 2005). States were required to set a proficiency level which can be based on the scores of the state's lowest-achieving demographic or the scores of its lowest-achieving schools, whichever is higher, as well as other indicators. In Illinois, for example, the 2002 assessment results were used as the baseline (Christie, 2003). Once the proficiency level has been set, Annual Yearly Progress (AYP), set by the state, must be achieved each

year. Annual yearly progress is a term to describe the incremental increase in scores that a school must meet each year. In 2002, the law's expectation was that students would eventually score at the 100% level by the years 2014-2015 (ASCD, 2005).

If a school has not met AYP two years in a row, the school receives technical assistance to develop a plan that will help the school get back on track. This plan might include receiving federal funds to be used toward school improvement. Schools that are designated as needing improvement would also be required to allow transfers for their children who choose to go to a better-performing school. Not only do they offer transfers, but students who have low test scores are eligible for money to pay for private or religiously affiliated summer programs that would offer additional educational opportunities. Once the school has met AYP for two consecutive years, the school is taken off the needs improvement list. However, If the school or district still cannot make the progress expected after four years under the Needs Improvement label; reconstitution, hiring of a private management contractor, conversion to a charter school, or staff restructuring are all possible consequences (ASCD, 2005).

By the time NCLB was passed, Tennessee had already been testing their students for several years. Therefore, Tennessee did not have as hard a time complying with the new dictates of the legislation as other states, which had either fledgling test systems in place or none at all. Tennessee developed an achievement testing system in the early 80's and began testing its students in 1983. Because the study was done in a county in Tennessee, it is important to understand the type of test the students are taking and the basic elements of Tennessee's Comprehensive Assessment Program (TCAP).

Tennessee's Accountability Testing Program

The Tennessee Comprehensive Assessment Program (TCAP, pronounced TeeCap) Achievement Test is part of the Tennessee Comprehensive Assessment Program (TCAP) that oversees all assessments offered by Tennessee. The TCAP Achievement Test is a multiple-choice test mandated for all students in Tennessee public schools in grades 3-8. The tests cover Reading/Language Arts, Math, Science, and Social Studies. In grades 5 and 8, a writing assessment is also given where students have 35 minutes to write an essay on a given prompt. Tests are taken over a five-day period and testing is done in the morning. Student testing time for each day ranges from 136 minutes to 104 minutes, depending on the section being tested. Tests are written in English only, no alternative languages are offered. The test is a version of McGraw-Hill's Terra Nova exam, specialized for Tennessee's curriculum (Tennessee Department of Education, 2010).

In addition to the achievement tests for 3-8th grades, the TCAP program includes an optional K-2 test that is norm-referenced, which compares those students who take the exam to the rest of their same age peers nationwide. For high school, Tennessee uses the Gateway, or end of course (EOC) exams in algebra I, math foundation, English I, English II, biology I, physical science, US history, and writing upon completion of relevant courses. Students must pass the algebra I, English II, and biology I tests in order to graduate. The EOC exam counts as 15% of the student's grade. All of the tests are standards-based, which means they measure specific skills defined for each grade by the state of Tennessee.

The TCAP Achievement Tests are criterion-referenced tests. Criterion-referenced tests are those that measure students' achievement against specific standards instead of measuring their performance against other students' scores. The test results show the level of proficiency reached for each subject. The scores are based on a Performance level designated by The Reporting Categories Performance Index, an estimate of the number of items the student would be expected to answer correctly if 100 of these particular items for the category were on the test. For example, there are 7 reporting categories for the Reading/Language Arts Test: Content; Meaning; Vocabulary; Writing/Organization; Writing/Process; Grammar/Conventions; Techniques and Skills. Each category has its own proficiency rating. The proficiency levels are Advanced, Proficient, and Below Proficient.

In addition to achievement data, Tennessee uses a 'value added' system that was developed by a professor at the University of Tennessee, Dr. William Sanders. The system, called the Tennessee Value Added Achievement System (TVAAS) tracks students scores over time (Tennessee Department of Education, 2010). TVAAS can measure a teacher's, school's or district's academic progress using achievement scores from students or groups of students from year to year. It can also be used to predict the growth that a student should accomplish based on past growth measurements. TVAAS adds to achievement data by allowing school's or districts to see growth over time. "By measuring the academic achievement of students and the academic growth of students, schools, and teachers will have a more comprehensive picture of the effectiveness in raising student proficiency" (McClure, 2009, p. 11). TVAAS scores are expressed in Normal Curve Equivalents (NCE). A TVAAS report compares the NCE to the growth

standard, the least acceptable measure of growth for one year of schooling and is shown on a scale in order to see a growth pattern for a student. This pattern is determined by using all previous existing student achievement data (McClure, 2009).

According to the Supervisor of Research and Evaluation in the school district in which the study was conducted, the TCAP Achievement Test data is used first and foremost for accountability purposes. Secondly, all of the schools receive a data notebook, enabling a school to see its yearly progress. The data notebook would also contain Tennessee Value-Added Assessment System (TVAAS) growth data to help schools see progress over time. This longitudinal data is helpful for program evaluation. Finally, achievement scores are also used when students change level (e.g., elementary to middle school) for placement (J. Beckett, personal communication, January 19, 2010).

While TVAAS is helpful for teachers, schools, and districts to view progress of individual students for Individual Education Plans (IEPs), parent conferences, discipline meetings, and targeted interventions; TVAAS scores are not used to as part of federal accountability requirements (McClure, 2009). The high stakes are tied to the Annual Yearly Progress; thus sanctions are given using only year-to-year data.

As stated before, education reform using accountability measures makes several assumptions about the effects of testing that involve the teacher including the assumption that instruction will improve, that tests will measure necessary outcomes of the school, and that tests will measure school/class/student success or failure (Madaus and Clarke, 2001). At the heart of these expectations, the main actors in this accountability movement are students and teachers.

Teachers' Beliefs about Effects of Tests on Their Students

Teachers are certainly the "ground troops" for high stakes testing reform (Crisp, , Williams,, Greenberg, , 2006). Clearly, in any new policy that affects the classroom, teachers are the main actors in the innovation. As such, research has shown they can influence outcomes of an innovation (Fuller, 1969, Hall et al., 1977; Richardson, 1990). In an article about teachers' responses to educational innovation, van den Berg (1999) makes a case for an increase in the attention to teachers' concerns. As van den Berg points out, "The specific reactions to an innovation and the possible problems associated with these reactions typically stem from the significance or meanings that teachers attach to their situation" (van den Berg and Ros, p. 881, 1999). Gitlin and Margoris (1995) also point out in their article about teacher resistance to change, that reforms do not consider important aspects such as teachers' perceptions of their role, teacher's interactions with one another, and how much the change will affect the teacher.

Teachers' attitudes, values and beliefs play an important role in a student's experience. A review was done of all studies that interviewed or surveyed teachers about the effects of high stakes testing. While many of these studies included effects that were specific to teachers, for the purposes of this dissertation, the studies were reviewed for specific information on how teachers thought students were being affected. Also, any effect the teacher mentioned that would certainly be experienced by students was included. Therefore, the review will only include information from these studies that affected students and will not include all effects observed by teachers.

A growing body of research regarding teacher beliefs of high stakes testing suggests that teachers have real concerns about the effects of high stakes testing on their students (Pedulla, et al., 2003, Jones, et al., 1999, Madaus, 2001; Nolen, Haladyna, & Haas, 1992; Darling-Hammond, 1991; Corbett & Wilson, 1989). Twenty-six studies were reviewed for this dissertation. Most of the research done has been survey research, but several studies interviewed the teachers and three studies were reflective exercises, where the teachers wrote about their own experiences under high stakes testing accountability.

Teacher's surveyed and interviewed reported on two main issues altering student experiences in school: the impact on classroom practices and the pressure of high stakes testing. The pressure of high stakes testing was found to impact both the morale and stress levels of students, as well as have effects on learning (Barksdale-Ladd, et al, 2000, Gordon & Reese, 1997; Hargrove, et al., 2000; Jones, et al., 1996; Darling-Hammond, 1991).

The Pressure of High Stakes Testing

One of the reasons I decided to interview students was based on results of an earlier study focused on revealing the lived experiences of teachers in a high stakes testing environment (Crisp, Williams, Greenberg, , 2006). When interviewing teachers about their experiences with high stakes tests, the number of statements made about the impact on students was revealing. Four themes came out of this study. The first theme, "we're the ground troops" is a metaphor on the teachers' role in educational hierarchy and cannot be overstated: any decisions made involving education reform inevitably settles

on the shoulders of the teachers to implement. The second part of this theme was “It kills you”, referring to a teacher’s feeling of watching students who struggle in a high stakes testing environment. The concern that teachers had for the students was revealed in other themes, such as the one titled “Motivation and Morale.” A quote pulled from this theme: “What’s sad is it doesn’t work for the kids.” These themes are evidence that teachers are concerned the effects of No Child Left Behind is negatively impacting the very students the law is trying to help (Crisp, Williams, and Greenberg, 2006).

What also stood out in the research was the strong language they used to describe the tests: “horrible”, “demeaning”, and “degrading” were all words used by the teachers interviewed to describe these tests. The words were not only strongly negative but directly tied to their great concern for the students. Quotes reflecting this concern were seen in statements throughout the interviews: “I just think it’s too much for our kids,” and “This test is everything. This hinges on whether you’re intelligent and smart” (Crisp, , Williams, Greenberg, 2006).

In a review of the literature, many variations of these themes were repeated in studies that interviewed teachers in high stakes testing environments. Gordon and Reese (1997) who interviewed 100 teachers in Texas, one of the first states to adopt high stakes testing discovered that teachers talk about testing’s effects on students in three ways: emotionally, academically, and socially. The emotional responses varied from no reaction to the tests to stress-laden reaction that “leads to anger and resentment” (Gordon and Reese, 1997, p.356). The teachers reported that there were students who had little or no reaction to the test, those who experience moderate stress This group of students were described as high achievers and the stress they experienced was caused

by the desire to do well on the tests. The last group of students were those that experienced a high level of stress (Gordon and Reese, 1997).

Another study surveyed teachers in North Carolina about the impact of high stakes testing on students. The percentages of teachers who felt their students were more prepared for learning (28%) and had more confidence (15%) was underwhelming considering that 61% reported an increase in their students' anxiety levels and 48.5% felt that high stakes tests had a negative impact on the students' love of learning (Jones, et. al., 1999).

Terzian, a teacher who chronicled her own experience teaching in a school that was labeled in the Needs Improvement category observed nervous behaviors from her students. One child in her classroom was physically shaking from nerves while another played nervously with his shoelaces. One expressed her anxiety verbal by exclaiming, "I'm going to fail!" when the test was placed on her desk. Terzian also perceived that her stress must be rubbing off on the kids, "If I was feeling this pressure, I could only imagine what my students were feeling" (Terzian, 2002, p.2).

Changes to Curriculum

Teacher's perceived stress was also connected to changes in curriculum. Jones and Egley (2007), who surveyed 708 3rd, 4th, and 5th grade teachers in the Florida school system found that those teachers who reported feeling the most pressure, were also more likely to spend more time on test-taking practice strategies in math, reading, and writing. Pedulla (2003) also found evidence those teachers in states who had high stakes spent more time on test preparation. Teachers also felt that the Florida Comprehensive

Achievement Test (FCAT) had a negative effect on learning and that students would have more knowledge and skill if the FCAT did not exist (Jones and Egley, 2007). In a study of 300 school districts in two states, teachers reported negative effects on teaching and learning when both stakes and pressure to raise test scores were high (Corbett and Wilson, 1991),

One of the biggest changes to curriculum after NCLB passed was the increased focus on reading instruction (Jones & Egley, 2007, Firestone et.al, 2002, Jones, Jones & Hargrove, 2003, Cimbricz, 2002). Schools across the nation have adopted programs that were proven to improve reading ability. While reading programs such as Success For All and Open Court have been scientifically proven to work (Slavin, 2006), some have criticized these programs as being too scripted and not having enough flexibility to meet individual children's needs (Alvarez and Corn, 2008).

One study asked teachers about their experience with the reading curriculum Success For All, and teachers did share that they had concerns about their students' learning. Success for All is a reading program that uses scripts, or set lesson plans that teachers use and follow exactly. The Success for All Program was designed to have all students reading at grade level by 3rd grade. One complaint was that the scriptedness of the program hinders a teacher's ability to make decisions independent of the curriculum in order to meet children's needs. Teachers were also concerned that the books were boring and that the redundancy of the program day after day turned kids away from reading instead of on to it. "And actually that is probably my biggest beef with SFA is that my students do not like to read anymore. Even the high ones do not like to read anymore" (Evans, 2008, p. 17).

Another study echoed this frustration of curriculum changes in reading due to high stakes, “I felt compelled to use teaching strategies that didn’t meet many of my students’ needs because of the high stakes attached to this assessment” (Alvarez and Corn, 2008). SFA is a program used in many schools that have at-risk populations (Slavin, 2008), therefore it has had a larger impact on those student’s most at-risk.

Concerns for Those Most At-Risk

At-risk students were of particular concern to teachers when looking at testing’s effects (Gordon and Reese, 1997, Alvarez and Corn, 2008, Firestone et.al., 2000, Johnson & Johnson, 2002). Johnson and Johnson (2002) spent a year in a Louisiana school, teaching and observing effects on teachers and their students under a high stakes testing system. In their book, they portrayed the grinding effects of poverty that affected all parts of the student’s life, including academic achievement. Added to this mix was the fact that demands placed on teachers under a high stakes accountability system stifled creativity and motivation for teaching. Teachers were extremely challenged to meet the needs of students and the needs of an accountability system that could shut down their school the following year (Johnson & Johnson, 2002).

Firestone et.al. (2002) surveyed 300 teachers in New Jersey and interviewed 30 of those surveyed. The results showed more direct instruction in poorer schools than in richer districts. Lomax et. al (1995), interested in effects of high stakes tests on minorities in Texas, found that teachers who worked in a school with a high minority population were more likely to teach in ways that were influenced by the standardized tests (Lomax, et al, 1995). According to teachers, students most likely to be affected by a narrowing of

instruction, drill and practice teaching methods, and scripted reading programs are those in lower in the districts with minority populations.

Thus, teachers have reported a variety of beliefs of the effects of high stakes testing on students. Teachers see an increase in stress in students. Effects on learning have also been reported, specifically in areas serving the poor and minority populations. A variety of changes in curriculum and teaching methods and programs are believed to negatively impact learning. In order to see if student's response reflects these beliefs, a review of children's perceptions of high stakes tests was done.

Childrens' Response to Tests

Although teachers' perceptions of high stakes testing reform has been the focus of almost 30 studies, there are only a handful of studies that look at students' perceptions. Three of the studies interviewed high school students, one study talked to both middle and high school students, one study interviewed only middle school students, one study involved 4th, 8th and 10th grade students and one study involved kids in 3rd through 6th grades.

Researchers that have done studies focused on student experience found that children were insightful and allowed the researcher a better understanding of the meaning of reform to students (Wilson and Corbett, 2001; Nieto, 1994; Ferguson, 2001). Wilson and Corbett (2001), the authors of *Listening to Urban Kids*, found that not only were children "useful windows through which to view reform," they also described them as "sophisticated observers of school life." Nieto (1994), who also calls for the use of student

voices for reform, argues that even though minority students who find school a waste of time can describe exactly how this perception was developed (Nieto, 1994).

Survey research has shown that students are much more positive about testing in elementary school than in adolescence (Paris, et. al. 1991, Debard, 2000). However, surveying elementary school children can be problematic. For example, in the Debard study, most fourth graders felt that testing improved critical thinking skills, and 8 out of 10 students believed that testing helped them apply information instead of just recalling it. It is doubtful that a fourth grader would have an adult educator's understanding of terms like "critical thinking" or "apply", and "recall."

One study that used student interviews for their data was focused on high stakes math preparation for sophomores. Six African-American students were interviewed at a large inner-city school in Ohio with a large minority population. All had good grades but did not pass the math portion of the Ohio Proficiency Test. The test was considered high stakes because these students must pass the proficiency tests in order to graduate (Lattimore, 2005).

Interestingly, in the Lattimore study, students did not express anger about the test; they accepted it and did not challenge its premise or applications in their responses. However, students did express frustration about the test preparation. "You cannot study and think over it the way you want to. You are trying to cram it, and the information you are cramming in is not clear." Lattimore concludes that African-American students preparing for high stakes tests are being exposed to what he coins a "pedagogy of mediocrity" (Lattimore, 2005, 144).

The pedagogy places restrictions on teaching practices that focus on computations

and procedures that are myopic and prevent the development of the essential conceptual understanding needed to navigate high-stakes tests (Lattimore, 2005, 144).

Wheelock, Babell, and Haney (2000) collected drawings to understand the perceptions of older elementary, middle school, and high school students on high stakes tests in Massachusetts. A research group made up of the researchers and educators who submitted the drawings for the study reviewed and coded the drawings to ensure internal validity. Some pictures, however, were fairly clear: In one picture, a student drew a picture of himself taking the test. Beside the drawing of this student was a thought bubble where he was burning the test with a lighter (Wheelock, Babell, and Haney, 2000).

The varied responses of the students were indicative of the complexity of student response to testing. While emotions such as anger, anxiety, boredom, motivation, and confidence were observed, a review of the percentage of drawings that portrayed each of these emotions showed there was no one overwhelming collective response to the test, nor did any themes emerge from the drawings that were consistent with all or even most of the drawings that could hint at a common experience.

Eighteen percent of the students drew themselves as “diligent and motivated test takers (Wheelock, Bebell, and Haney, p. 15, 2002). Out of these pictures a sense of confidence was noted in the drawings. Seventeen percent of the students’ drawings indicated difficulty. One out of 12 of the drawings showing difficulty indicated that the content was confusing. Twenty-three of the students’ drawings showed negative feelings: 13% of the images portrayed students as anxious and 10% of the students’ drew pictures depicting anger towards the MCAS (Wheelock, Babell, and Haney, 2002).

Triplett and Barksdale (2003) also were interested about children's thoughts and feelings. In a previous study, where they interviewed young children about general testing experiences, they found that 17 out of the 33 participants experience worry and anxiety specifically over the Florida Comprehensive Assessment Test (FCAT), the high stakes test given to Florida students (Triplett & Barksdale, 2003). These students talked about having physical symptoms that are normally connected to stress, including stomach aches and loss of sleep (Triplett & Barksdale, 2003).

Triplett & Barskdale (2005) later collected drawings of 225 students from five different schools in two different states. Their participants were in 3rd-6th grades who had just taken a high stakes test much like the Wheelock and Bebell study. However, in addition to asking students to draw a picture, they also asked the students to write a description of their picture.

Triplett and Barksdale (2005) identified nine categories from the analysis of both drawings and written descriptions: 1) Emotions; 2) Easy; 3) Content Areas; 4) Teacher Role; 5) Student Metaphors; 6) Fire; 7) Power/Politics; 8) Adult Language; 9) Culture of Testing.

The researchers found a prevailing negativity among the 225 pictures and the descriptions that students wrote, explaining their pictures. The category "Emotions" was the most prevalent category that came out of the study. "The recurring words that expressed emotions included: 'nervous,' 'mad,' 'sad,' 'frustrated,' 'hate,' 'confused,' 'bored,' 'tired,' 'sweating,' and 'sleepy,'" Out of all of these emotions "nervous" was the one reported most often (Triplett and Barksdale, 2005, p.244).

Easy was the second most prevalent category but had mixed responses. While kids said the word easy in their written response, they would also say that sometimes it was confusing or hard. Twenty-eight students drew and wrote about the content areas of the test. For the 19 students whose picture or writing depicted a teacher role, it was either as a monitor, coach, comforter, or uninterested observer. In the Student Metaphor category, one student write about what was going on in his head, "I felt like there was a war going on in my head. The light bulbs won and the question marks lost! (Triplett and Barksdale, 2005, p. 249)" Among the eleven identified categories, the category of Fire emerged as many children drew the test or school in flames. There was a similar pattern noticed in the Wheelock et. al. study.

The category Power and Politics contained drawings and descriptions that mentioned testing as serving a political function or the power that the test wields. In the Adult Language category included drawings that literally contained curse words or that portrayed adults talking to the students. If the drawing included the aspects of taking tests, such as filling in bubble sheets, then it was added to the Culture of Testing category (Triplett and Barksdale, 2005).

Roderick and Engel (2001) conducted a study to see if a reform initiative based on a high stakes test and increased resources would improve effort among minority students whose test scores showed they risked being retained. Interviews from 102 students in 6th and 8th grades revealed that there were motivational increases among students who were categorized as having "high levels of work effort." This group of students also showed gains in both learning and positive promotional outcomes. However, results showed that there was a significant group that did not increase effort despite a desire for promotion.

These kids were not engaged due to a lack of motivation stemming from barriers to learning due to home, behavioral, or learning issues (Roderick & Engel, 2001).

Another finding from this mixed-method study suggests that age was a factor in understanding certain important learning tools for success. For example, findings suggest eighth graders are more likely to understand that improving skills is related to a better outcome on the test. The older students were also able to understand learning concepts that are seen in successful students, including the idea that passing the test is an important goal to be achieved. The students were also able to connect their increase in working harder to a change in attitude and behavior (Roderick & Engel, 2001).

Another study with sophomore students was conducted in a single school in Indiana where they too had to pass a test in order to graduate. The students, who were given surveys and interviewed had some interesting – but at the same time - contradictory things to say. Half of the students were afraid they wouldn't graduate, however, they were only minimally anxious and 89% of them felt that their parents weren't worried. This information surprised the researchers (who were teachers at the school). According to previous test scores of these students, many of these students were at great risk for failure.

In the follow-up interviews the students shared what the test meant to them which helped explain some of their responses. Students explained that they weren't worried about passing the test. To not graduate was not the end of the world. One student replied, "If I fail, I fail. I'll just get my GED" (Hughes & Bailey, 2002, p.75). Others interviewed reasoned that the test was unfair, therefore it was not worth getting upset about. Some explained that they knew students who did well in school, but who didn't do well on the

test, thus the results were not reflective of their abilities. Researchers heard the phrase: "It doesn't prove anything" repeatedly in their interviews (Hughes & Bailey, 2002, p.76).

Conclusions

Since 1965, efforts to both improve and equalize the educational system have used assessments to look at results of federally funded educational programs. The methods of testing have evolved as each reauthorization has attempted to fix the problems of the previous ESEA laws. No Child Left Behind, the last reauthorization of ESEA, increased the federal role in education by requiring end of the year assessments for all states and tied sanctions to the results of these tests if schools haven't met the annual yearly progress goals set by the state.

Teachers in states such as Florida, North Carolina, Florida, Texas, and Tennessee have been surveyed and interviewed. Concerns were documented about accountability testing and its impact on students, including an increased amount of pressure on students and detrimental changes to curriculum. Concerns for students who are most at risk including children in low socio-economic brackets and minorities were also voiced. Teachers' concerns affirm the need for studying student's reactions to these tests.

In previous studies done exploring student beliefs and attitudes about high stakes testing, a range of feelings were expressed. Each individual study revealed important findings In the studies done using student drawings for their data, both found a sense of anger and anxiety among the participants. However, the Wheelock, Babell, and Haney (2000) study revealed more positive responses than the Triplett and Barksdale study

(2005). Triplett and Barksdale found that students react negatively to certain aspects of testing culture including the timed tests and the long periods of quiet. These studies both showed that students had definite emotions tied to the tests that have potential to affect student performance and motivation. Wheelock, Babell, and Haney (2000) concluded that children's reactions to high stakes testing refute claims made that the MCAS will improve student motivation.

The Roderick and Engel (2001) and Hughes and Bailey (2002) studies revealed the potential of interviewing students to understand the meaning of high stakes tests. For example, Roderick and Engels' (2001) findings show that student's attitudes and beliefs in learning have a positive effect on test scores and are more likely to develop in 8th graders rather than in 6th. When interviewing high school students, Hughes and Bailey (2002) found that interviewing students revealed the meanings of the test to students which helped explain their survey responses that had at first confused the teachers at the school.

In conclusion, research using student data reveals important findings with regards to students taking an accountability test. From the studies done, it is clear that students have an array of emotional reactions to testing that could have major implications on the student performance and thus the results of these tests. According to the research reviewed with older students, the meanings of both middle and high school students' affect can be revealed if we interview them, adding important information that cannot at first be completely understood with surveys. The assumption is that 4th and 5th graders are also reliable reporters of their lives and, when interviewed, can help us to understand student's experience of a high stakes test in younger students.

Chapter III

Introduction

In chapter 3, I will discuss why the phenomenological method was used in this study. Next, I will explain the steps of the phenomenological research method. Finally, I will describe how this particular phenomenological study -exploring the experience of students taking a high stakes test - was conducted.

Why Phenomenology?

A criticism of education as a discipline is the inability to divvy up the child into analyzable elements as well as to analyze an important element of education: the relationship between the child and teacher (van den Burg, 1974). Hirst implicated a solution to this problem in the discovery of the "realm" of education and is paraphrased by van den Burg (1974) in the following paragraph:

> Hirst claims that conceiving of educational theory as educational principles connected in one direction to the more abstract, special sciences and in the other direction to the more concrete situations of practice, but in either case connected merely by an open-textured, informal, contextual logic, creates an autonomous realm, a real space of province of meaning, in which to place educational theory. This realm in turn, allows the concrete situation of educational practice to emerge in its wholeness, unreduced by the methods of inquiry of the special sciences. It lets education-and educating-appear. This is the prerequisite for the development of the study of education as a discipline in its own right (p.185).

Phenomenology offers a method to see Hirst's realm described above, to look at something as a whole; a method that can allow for a context and meaning. Polkinghorne (1989) explains that phenomenological research "reminds us that the research journey needs to attend to the configurations of experience before moving on to assumptions" (p. 41). Therefore, phenomenology is not theory based, nor is its intention to produce theory. However, like all research it makes certain assumptions. For example, phenomenology believes we can 'bracket' or set aside our assumptions of a phenomenon studied by sharing these assumptions at the beginning of the research process. Phenomenology as a research method also believes that we can understand as well as reveal the participant's experience of a phenomenon by using the words of the participants as data. Phenomenological research produces a description that captures the essence of a situation that has meaning to the participants in that study.

Phenomenology is grounded in Continental philosophy and is intended to shed light on the lived experience of research participants. Developed by Husserl and later expounded upon by Heidegger and others, phenomenology was a method that existentialists began turning to when probing the depths of who we are and the path to living an authentic life (Pollio, Henley, and Thompson, 1997). Fischer (2006) defines phenomenology as "the study of phenomenal accounts [self-reports] to grasp what they imply about lived relations to the object or situation; about assumptions, meanings, past and future as well as present; and the relations of discerned aspects of a phenomenon to each other and to the overall understanding" (p. 413). Thus phenomenology is a method focusing on the lived experience (Polkinghorne, 1989).

While many methods of phenomenology exist, they are influenced by three schools of phenomenology: the Duquesne school influenced by Husserl, the Dutch school influenced by Husserl and Heidegger, and the method used by the University of Tennessee, developed by Howard Pollio (Pollio, Graves., and Arfken, 2005, Thomas and Pollio, 2002, Pollio, Henley, and Thompson, 1997). The latter method is influenced not only by Husserl and Heidegger but also draws from the philosopher Merleau-Ponty. Merleau-Ponty's explication of intentionality and perception are key concepts influencing the UT school (Pollio, Henley, and Thompson, 1997).

The importance of Merleau Ponty's conceptual understanding of intentionality and perception in phenomenology is that it does no less than shape our experience and define who we are. Perception is the process whereby sensory input is transmuted into organized experience through the interaction between the person and the world. If perception is where we connect to the world then intentionality is how we connect to the world. As we navigate the world through our perception, we develop intentionality. Intentionality is the fundamental structure or pattern of the human experience (Pollio, Henley, and Thompson, 1997). Thus, Merleau Ponty reshapes the definition of "being in the world". This point is illustrated by Merleau-Ponty: "In the action of the hand while it is raised toward an object is contained a reference to the object, not as an object represented, but as that highly specific thing toward which we project ourselves, near which we are, in anticipation, and which we hunt. Consciousness is being toward the thing through the intermediary of the body" (pp. 138–139).

Experts point out the need for research that can help us understand the role of affect in learning (Nias, 1996; van den Berg, 1999). Prior phenomenological research has

suggested that the lived experience of children “appear to be unique and cannot be deduced simply by subtraction from the more ‘developed’ experiences of adults” (Briod, 1989, 124). McPhail (1995), who conducted phenomenological studies exploring adolescence experience found that the focus on human consciousness can influence our research and practices. Phenomenology also increases our understanding of the socio-cultural aspects of learning that are so important (McPhail, 1995). Because this study investigates the experience of students taking a high stakes test, the phenomenological method was chosen as the best way to get at a child’s experience.

The Phenomenological Research Method

For more than five years, the Applied Educational Psychology (AEP) research team, an educational research team at the University of Tennessee, has been using phenomenology to explore the lives of teachers and students in contexts such as working with underachievers, experiencing the world of high stakes testing, and experiences of a new program at a student athlete support facility. In the book *Listening to Patients*, the authors found that studies in phenomenological research on aspects of nursing went beyond the limits of survey research to uncover new aspects of the participants’ experience (Thomas and Pollio, 2002). The authors also found that phenomenology gets to what is most meaningful or prominent to the participants’ understanding of an experience. Discoveries in the past few years made by the Applied Educational Psychology Research Team have also made such observations.

The phenomenological method is marked by context centered, collaborative, communicative, and open-ended inquiry. The methodology is described in two edited

books, one in the field of nursing (Thomas & Pollio, 2002). This approach employs unstructured open-ended interviews. The researcher defines the starting question for interviews, but it is the participants' perception of their experience that directs the course and the content of the interview. Interviews are audiotaped, transcribed and analyzed in a phenomenological research group. These steps were taken verbatim from a paper presented to American Educational Research Association Annual Meeting in 2006 describing the UT AEP research team approach:

1. The opening question allows for a broad range of descriptive responses across participants. Researchers work with a research group to make sure the opening question will lead to a complete description of the lived experience of each participant interviewed.
2. Bracketing interview: A member of the research team interviews the researcher who will be conducting interviews—by asking the researcher to answer the same question as the participants will be asked, either based on personal experience or, if no personal experience with the focus of the research, then by sharing what he or she expects to hear from participants. The purpose of the bracketing interview is to help the interviewer become aware of historical and personal contexts and especially assumptions, so that the interviewer can avoid asking leading questions of the participant and become more open during the interview and later, during the analysis.
3. Before an interview begins, the participant is asked to read and sign a consent form that explains how confidentiality will be maintained, according to the UT human subjects mandates based on national standards. Then, the interviewer

makes an audio tape of the interview. The interview begins with the opening question. Additional questions are not determined in advance, but are asked as needed to help the participant clarify, refocus on unfolding themes, add details, and provide examples. The purpose of all interview questions is to seek as complete a description as possible of the lived experience being researched. Additional participants are interviewed until a repetition of themes begins to stand out, and no new themes are detected. Since the purpose of phenomenological research is to determine a clear picture of the essence of the human experience related to some phenomenon, a relatively small number of participants is needed. Multiple participants are included to improve the researchers' interpretive vision by considering more diverse experiences and finding similarities.

4. Interviews are transcribed, identifying information removed or changed, and accuracy double-checked.
5. In the UT Applied Educational Psychology research team approach, members of a research team interpret the data together, or at least major parts of the data. Each participant's transcript is read out loud while team members connect what stands out to them, which becomes a basis for themes. (Themes are experiential patterns exhibited in diverse situations.) Themes are determined first within each transcript separately, then by looking for commonalities across transcripts. In both situations, research team members challenge each other to find supporting quotes within transcripts for any interpretative theme. A theme identified within an individual transcript becomes a common them only when a majority of participants have expressed its meaning as belonging to the experience. Together, the team

seeks a consensus and labels themes using the words of the participants', seeking especially metaphors as labels whenever possible. (Metaphors are often used in descriptions when ordinary words fail to adequately express meaning. They reveal aspects of subjective matter otherwise ignored or undiscovered.) This method is based primarily on Gadamer's ideas regarding hermeneutic analysis (Gadamer, 1966).

6. A structure is determined for common themes that include a contextual theme representing all participants from which the other themes appear. In addition, our teaching/learning research team utilizes a method described by Polkinghorne for representing the essence of the lived experience related to the phenomenon under study. Instead of stating the essence in abstract language recommended by Polkinghorne, our research team uses primarily the words of our participants that represent the themes—as though the participants were speaking in one voice.
7. The UT Applied Educational Psychology research team method also involves two procedures for seeking reliability (whether someone else would derive similar themes; i.e., thematic consistency) and validity (whether convincing evidence exists for descriptions/quotes identified from the transcripts). The first procedure involves the researcher or research team presenting themes to an interdisciplinary research group, with supporting quotes. The external research group determines whether they believe evidence exists that supports the themes. On some occasions, the external group actually analyzes a transcript or two from the study and the research team compares the results to their own interpretation, possibly adding or deleting themes. The second procedure involves submitting the results

to participants to seek their reactions, resulting in further analysis if participants do not agree with the results.

(Crisp, Williams, and Greenberg, 2006)

Bracketing Interviews

Two bracketing interviews were done to explore my assumptions about the phenomenon in the presence of the participants of the Applied Educational Psychology Research Team. A bracketing interview is conducted by one or more of the research team and interviews the main researcher on their assumptions about the phenomenon about to be explored. One was done prior to my first round of interviews that were not included in this research study. This pilot study focused on fourth grade students who did not do well on the TCAP Achievement Test based on the previous year's scores. This study recruited students in an inner city school in Knoxville and was intended to be the original dissertation study. However, the students, despite my efforts at befriending them and making sure they were comfortable, did not give me enough data for themes to develop. I used this experience as a learning opportunity in proper interview technique. I also felt that students would be more comfortable talking about their experiences outside of a school environment, where they would perceive me more as a student and not as a teacher or authority figure.

In this second bracketing interview, based on my previous experience, I held the assumption that students would say the tests were hard, and that they would talk about

the length of the tests. It is important to note that in my first round of interviews, I chose students who did not do well on the TCAP Achievement Test. Also, I also assumed that the students may be worried about the test.

In my second bracketing interview, the focus was on exploring my assumptions now that I had already interviewed a set of students. I assumed that the students may tell me what I described as cover stories, or would tell me what they thought I wanted to hear. I talked about how, in my last interviews, students would sometimes talk into the microphone like they were reading a script or answering a question that they wanted to get right. I held the assumption that the indoctrination the students receive about the test would cloud the thoughts and feelings of the students. However, one of the committee members astutely said that that the teachers' role is part of their experience and that possibly, wanting to say the right thing or do their best was the meaning of their experience of this test.

Also, by this time, my son was two months into a school year at a school that failed to meet Annual Yearly Progress for the second year and was put on the School Improvement I List. The different levels of high priority schools/systems are: School Improvement 1, School Improvement 2, Corrective Action, Restructuring 1, Restructuring 2 and SEA/LEA Reconstitution Plan. I was much more aware of the negative effects on instruction and time management of the school than I had been prior to the last study. Based on my experience in this school, I felt that the students would feel over assessed, as they have to take weekly TCAP-style assessments, as well as periodical benchmark tests, all in the multiple choice, standardized test format. Playtime was shortened and science was not taught at all the first six6 weeks of school. Acknowledging that I was

personally affected by the very phenomenon that I was studying was certainly an important thing for the entire research team to understand when we later analyzed the transcripts together. It was also revealed that I assumed kids would not mention parents in their interviews, as this had been my experience with the previous round of interviews.

Identifying the Participants

Students from a recreation center were recruited to participate in the study. The recreation center is a free afterschool program that caters to a community within the city of Knoxville. Students had the opportunity to play basketball in the gym, play tag outside on the lawn, or play games in the game room. Students were given a snack when they got off the bus from school and then they play a structured game together, usually involving lots of exercise. The rest of the time was spent in free play. Children are picked up by their parents or walked home by 6 pm, when the center closes.

The center was a positive afterschool environment. The director ensured the kids played safely and attended to any disciplinary issues. A respected leader in the center, the kids liked him and wanted to please – at least most of the time. He focused on building character traits when students made the wrong choices. Often, students were given a chore to do to make things right if the child has done something wrong. He constantly circulated among the rooms and was a constant steady mentor, especially to the boys at the center.

Students from both schools attended this popular afterschool program. Both schools were classified Title 1 schools, a designation received when at least 40% of the

population receives free or reduced lunch at the school. Both schools well exceeded the required percentage. Eighty percent of School 1's student population was economically disadvantaged and 85% of the other school's population had this label (Tennessee Department of Education, 2007). At the time of the study, School 1 was classified in the School Improvement I category based on the previous year's TCAP Achievement Test scores (Tennessee Department of Education, 2007). By contrast, School 2, despite it being an inner city school and having a larger population of students with disabilities, won an award for making the largest gains on the TCAP Achievement Test in their school district in 2006.

Students were sent home with consent request forms. For two days, I handed out consent forms to the students at the center. About 40 students were handed forms on both days. Out of the students participating in the recreation center after school program, nine forms were returned. All interviewees were between the ages of 9-11. Pseudonyms were used for the student participants in the study to ensure confidentiality. Carrie and Sadie, two Caucasian twin girls, attended School I. Frank, an African American, was also from School I. The others; Kevin, Mark, Deshawn, Chris, Kevin, Olivia, and Kathy all attended School 2. Deshawn, Kevin, and Chris are African-American, Mark, Olivia and Kathy are Caucasian.

While I did not gather SES data, the kids who typically go to free afterschool programs at recreation centers are generally kids from low-income families. Since over 80% of the children in both of the schools receive free lunch, it is probable that at least a few and probably all of these students qualified.

I also did not gather data on how well the participants did on the TCAP Achievement Test. My objective was to find the essence of children's experience of taking the TCAP Achievement Test. It was not my objective to compare or contrast differences in experience between those who did well and those who did not do well.

Interviews

Hatch (1990) identifies potential problems when interviewing children. When adults interview children there is an obvious status difference that some children react to more than others. This status problem can lead to children wanting to give adults the right answer.

These were problems I encountered during my pilot study and I believe were exacerbated due to the school location of the interview (I found out after finishing my last interview in the previous study that the conference room we used was typically used for disciplinary issues). For this reason, I decided to gather data from a more neutral place. The recreation center, where children spend their time playing and having fun was a setting that I thought would get more relaxed students, who weren't worried about how they answered my question. Also, the adult supervision there was less authoritarian than in a school setting which I believed helped the participants open up more in the interviews.

Interviews were conducted at the recreation center in the office of the director of programs at the facility. It was a space that afforded privacy, but unfortunately, was not soundproof. Because the office was located right next to the gym, we were interrupted

during a couple of the interviews by basketballs hitting the window or an occasional child running by yelling loudly to a friend or teammate. Once we had to pause the interview because of a phone call.

The students recognized my presence in the recreation center community as 'Ethan's mom' who picked Ethan up every day. It was my habit to spend a few minutes playing basketball or talking with the director and a few kids who liked to tell me what they had been up to that afternoon. Some of the kids also attended Ethan's school and saw me there in the hallways volunteering each Friday. I did not, however, have a personal relationship with any of the kids interviewed. I was simply a casual acquaintance who they saw a few minutes each day.

Verbal assent forms were read to students informally and their verbal consent was recorded along with the entire interview. To further reduce the risk of them wanting to give me the right answer, I told them that I was really interested in *their* perspective. I told them that adults had a perspective about the TCAP Achievement Test and I already talked with them, but I was really interested in understanding what kids thought about taking the TCAP Achievement Test. Explaining that their view may be different than adults view helped them to understand that the kids view was unique and further reduced the chance of the participants giving me what they thought was the right answer or what I wanted to hear.

Interviews were recorded on a digital recorder. The interviews were unstructured and open-ended. Each participant was asked, "What was it like for you when you took the TCAP test?" Follow-up questions such as, "Can you say more about that?" were asked if a student mentioned a particular feeling or emotion where more information was needed

that might help the research group understand the meaning of the statement. At the end of the interview, the student was asked if they had anything more to say. If they said no, the student was thanked and the interview ended.

None of the interviews were more than 20 minutes long and most lasted 6-10 minutes. The interviews with children were much shorter than previous interviews with adults. The researcher transcribed each interview and all participants were given a pseudonym to protect their anonymity.

Interpretation of the Data

The Applied Educational Psychology research team conducted the interpretation of the data. Each member of the team signed a statement of confidentiality before we began reading the transcripts. The transcribed interview was read aloud in sections. We stopped when we felt we needed to go over an important element or if we had read over enough data that needed to be analyzed before continuing to read the next section. All of the interviews were read and discussed by the research group. The purpose of the discussion was to talk about what stood out to the group members in order to get the meaning of each portion of the manuscript, allowing other members to affirm or counter this impression.

The research group was an important part of the analysis of my study. It provided the rigor that is characteristic of a phenomenological study in two ways. One, the group, knowing my biases or assumptions from the earlier bracketing interview, could bring these up in discussion if at any point they felt there was a need to do so. Also, the

different backgrounds and perspectives of a group allows for a thorough interpretation of the data.

Then, on my own but with the group's input, I developed and organized quotes from transcripts into categories. I brought the categories back to the group for further analysis. Categories were challenged and others were proposed. To ensure an internal check for validity, all members had to agree on each category.

Themes then began to emerge as we continued to discuss and analyze the participant's words. Themes are patterns of descriptions that are recurrent across participants. When a theme was proposed, we reviewed the transcripts to ensure that data existed for each. We focused on, instead of identical detail across multiple participants, the overall essence of the lived experience as described by Valle and Halling (1989). Themes were not discussed with participants as a second source of outside validity. At the time, I felt that students may not understand the concepts of themes, and the objective of finding the meanings of experience would be hard to explain to children. However, after three years to reflect on it, I think it would have been an important step that had the potential to be illuminating.

In summary, phenomenology, marked by collaborative, communicative, and open-ended inquiry, illuminates the universal experience of the research participants. Influenced by Merleau-Ponty's definitions of perspective and intentionality, and the importance of context all influence this method. Finally, the methodology is rigorous, with measures put in place such as a research team to analyze data and the bracketing interview to reveal potential biases.

In an apt metaphor in *Listening to Patients* (2002), phenomenology is a lens, not a hammer. By describing values, meaning, intentions, morals, feelings, life experiences, and creations of human beings, phenomenology provides a conduit to understanding a student's being-in-the-world (McPhail, 1995). Through the use of rigorous methods and procedures, we can compare these findings with other high quality research. Chapter four will address the chapter findings of this data analysis process.

Chapter IV

Findings

The purpose of this study is to explore the experience of fourth and fifth grade students who took the Tennessee Comprehensive Assessment Program (TCAP) Achievement Test for grades 3-8, a criterion-referenced test given annually, primarily for accountability purposes. Nine students from two schools in East Tennessee were interviewed about their experience of taking the test. As outlined in Chapter III, the phenomenological method developed at the University of Tennessee by Dr. Howard Pollio was used to conduct the study (Pollio, Graves., and Arfken, 2005, Thomas and Pollio, 2002, Pollio, Henley, and Thompson, 1997).

The descriptions of the experience of students taking a high stakes test shed light on the four major grounds of human existence: others, world, body, and time. Merleau-Ponty (1962) and other existentialists postulate that these four grounds are universal to the human experience (Heidegger, 1962). The holistic aspect of phenomenology is important in understanding its significance; while phenomenology is concerned with the person, it does not only focus on thoughts and feelings, but an embodied “being-in-the-world” (Merleau-Ponty, 1962). This means that these students, reflecting on a test situation in which they find themselves, are also responding to this context. This response can be described in terms of existential grounds of the body, the mind, within relation to others, or within a framework of time.

The thematic structure of the experience of the participants taking a high stakes test includes four themes and a ground. These themes are as follows:

Ground: *It just felt like another test, but like more important/weird (Deshawn)*

Theme I: *It was kind of hard but kind of easy (Kathy)*

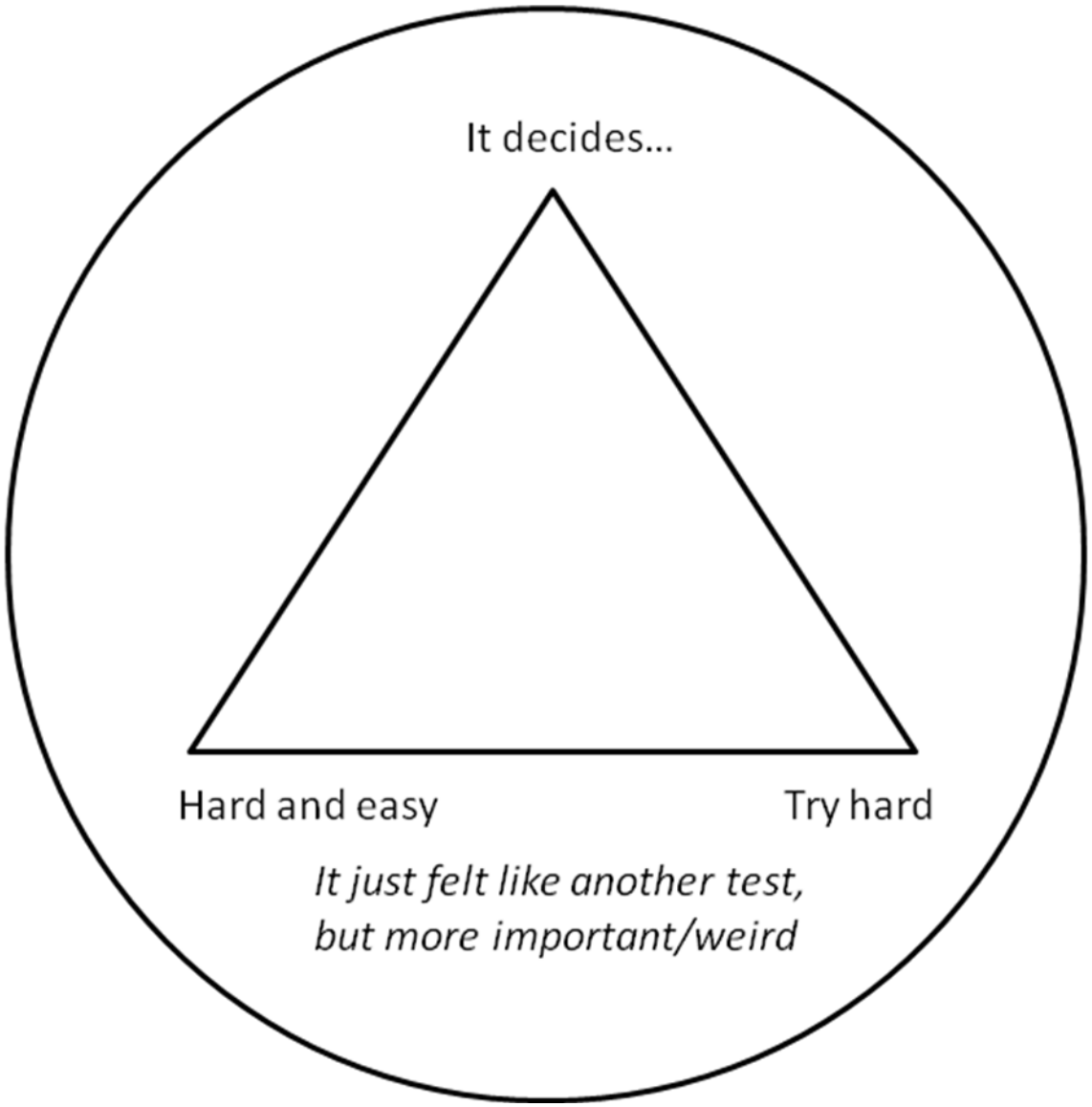
Theme II: *Yeah, you have to really try hard (Kevin)*

Theme III: *It decides, well, I don't know, but it seems… (Carrie)*

Figure 1 shows the structure of the experience of these students. Students experienced both the context and content of high stakes testing. The context or ground answers the question, "Where must I stand to see what you see." The taking of the TCAP Achievement Test is experienced by the child as like and unlike other tests. The test is important and weird because things happened during the week of the TCAP Achievement Test that did not occur in the regular school culture. The students understood that the test has hard and easy parts as the first theme indicates and the second theme describes the strategies that they used during the taking of the test. The third theme indicates that it's unclear of how the test is important even though they knew that the test is important.

The ground and three themes are interconnected. The ground, Theme 1 and Theme III seem to relate as the students understand the test. Theme III deals with the taking of the test as the student sees it. The themes are explained in detail on the following pages.

Figure 1: Structure of the Experience of Students Taking a High Stakes Test



The Ground

The quote representing the context comes from the words of Deshawn who *didn't really have trouble with [the test]* and Carrie:

It just felt like another test, but more important/weird (Deshawn and Carrie)

Students were grounded in their experience of tests they took in the past. In the interviews, when I asked them to tell me about the TCAP Achievement Test, they reported on the differences of the test – differences from *another test*. During the interviews students talked pragmatically about the test, talking about specific aspects of the test culture that separated it from a regular test.

This theme came from the reporting of the participants on the aspects of the test that made it stand out as different from a regular assessment. As the students described their experience, they pointed out the specialness of this test by talking about how the test altered their teacher's role and their teacher's physical presence.

However things about the experience with the test were not just important, but weird, "well, it's kind of *weird* every time." Weird is a word to describe something unusual or out of place. This comment and all others in the results section are grounded by the student's past experience of test-taking.

Carrie described a time in the test where she was unsure of a question and needed her teacher's help. The teacher's response was different than her ordinary response:

> *I raised my hand to ask my teacher. And she would say I can't help you with that. And I knew she couldn't but when she, I know what they mean when they can't help me because it would give away the answer, so when she said, I said o.k.*

In a standardized test, fidelity, or following closely to a set of standardized procedures when giving the test, is extremely important to assessment specialists as it affects the validity of the test. All proctors, usually teachers, have a checklist of items they must comply with that outlines specifically what is required to ensure all students receive a consistent environment for the test as well as protecting the security of the test. A Test

Administration manual is given to each school administrator who trains all of the proctors for each of the test that includes the checklist and there is a training session prior to the test to go over the details.

The checklist includes specific directions on the kind of assistance the test administer can provide: “Do not provide assistance that could indicate an answer” (Tennessee Department of Education, 2010, p. 19).

Kathy, like Carrie explained how her teacher interpreted this directive:

> *the teachers, they couldn’t help us; they could only read the question for us*. She goes on to explain, *they could only read the question for us cause like if they um, if they told us the answer, it would be cheating.*

Kathy and Carrie did not express strong negative or positive feelings about how they felt about their teachers not being able to help. Nor did they talk about how they felt about the idea that others might not trust their teachers. Students seemed matter-of-fact about the restrictions of the test, accepting this as a different kind of test. The girls understood that this was not the teacher being unhelpful to them, but part of the culture of this ‘more important’ kind of test.

Students were also aware of another way that teachers were kept from cheating that made this test stand out from other tests. Students spoke about their own teachers not being in their classroom in an effort to ensure that they wouldn’t help the students too much, or give students answers, what the students labeled “cheating”. Kathy provides another significant quote in this regard:

They could only read the question for us cause like if they um, if they told us the answer it would be cheating and that's why we had another teacher in our room too. And, um, because most of the classes they switched teachers because um, because they think that like [the teachers will] cheat and give you the answer. Our teacher told us that she said we was lucky because a lot teachers, they switch, but we just had another teacher (Kathy).

The students explained that not only can teachers not help answer the question, but the schools that these students attended also had the teachers change classrooms. This means that their regular teacher was not in the classroom with them during this important test.

Pressures of testing have also increased the amount of cheating on these high stakes tests (Yakimowski-Srebnmick, 2001). Numerous allegations of cheating have been documented in states across the country (Axtmen, 2005). Much of the Teachers Administration Guidelines is focused on providing checklists to ensure the security of the test, including discouraging the act of cheating. All teachers must also sign a Testing Code of Ethics that states the following:

> T.C.A. 49-1-607. Noncompliance with security guidelines for TCAP or successor test. — Any person found to have not followed security guidelines for administration of the TCAP test or a successor test, including making or distributing unauthorized copies of the test, altering a grade or answer sheet, providing copies of answers or test questions, or otherwise compromising the integrity of the testing process shall be placed on immediate suspension, and such actions will be grounds for dismissal, including dismissal of tenured employees.

Such actions shall be grounds for revocation of state license. I hereby certify that I have read and understand the State of Tennessee Test Security Policy.

Furthermore, I agree to abide by state test security guidelines and understand that any breach in test security on my part could lead to my dismissal and/or revocation of state license.

Teachers, by signing this document, acknowledge that they understand the high stakes of certain behavior in the distribution, implementation, and collection of the tests. In these two schools, the administrators have also discouraged the act of cheating by switching teachers to different classrooms.

Students clearly understood why their original teacher was not in the room. Evidence from student responses also revealed that teachers explained that they wanted to be in the classroom with their students. Mark explained,

Our teacher said she would if she could be there with us; she wants us to do good on it. She would help us if she could help us.

Certainly, having a different teacher in the room in elementary school was noticed.

Students revealed that that they were protective of their teachers and believed in their integrity, even though this policy seemed to be based on a suspicion that teachers would cheat. Carrie explains,

> *I had Miss _____ , so I might have to have Miss _____ to do the TCAP for me in the class and Miss _____ might do my sister' class, because um, otherwise your teacher might, but probably won't, but some teachers they might tell you the answer, so they have so we have to switch teachers.*

The students experienced their teachers as people who *probably wouldn't cheat* but just in case, the policy was put in place. Students were protective of their teachers and accepted their teachers' explanations of why they couldn't be there.

Being quiet was another aspect of the test that was different. Certainly, in the assessments students take during the regular school year, there is a modicum of quiet required, however, students pointed out that this was a different kind of quiet.

> *It's kind of weird every time, because you have to be so quiet and stuff, if you kind of get loud, it's kind of bad. But it's kind of good when you take a break, because it's um, you have to be quiet and stuff.*

Carrie used the adjective "weird" to describe many aspects of the test, from the length, to her reaction of having to be quiet during the test. She was reacting to this difference and found it to be strange or unusual in comparison to other tests. Carrie did not like "being" quiet, which for her is an action that she had to maintain during the test. She juxtaposes taking a break, when she did not have to expend this effort, as good.

While Carrie experienced being quiet as an action, Mark experienced the quiet in a different way, *it's just peace and quiet so...*. While both of these kids experienced the quiet during the TCAP Achievement Test, they attached different meanings to the quiet, one felt it was weird and something one has to maintain, and the other thought it was peaceful. However it was perceived, the quiet stood out as another aspect of an important test.

Another way students saw the test as important is in the snacks they received during the test:

> *We got snacks, and before the test we got like peppermints. And so I like that (Deshawn).*

Four of the students noted the novelty of a snack during this special test. All of the students that mentioned the snack thought it was a good thing. Kevin understood the meaning behind the snack,

> *The teacher allowed us to have snack and um, we had snack, we had juices and stuff and that helped us to concentrate more and sometimes on the TCAP we got to take peppermints and that also helps us concentrate (Kevin*).

Students are not used to getting snacks or peppermint candies during a normal test. These special treats that the students like were another indicator that this test was important.

Theme I

The name of Theme I came from Kathy who describes the reaction across participants of the content of the test:

It was kind of hard, but it was kind of easy (Kathy).

Students described the test as both hard and easy. In six of the interviews they used both adjectives or similar ones, to describe the test. In a phenomenological interview, asking back when participants use a word that describes a feeling is a good way to get more information about the significance of this word. This technique was especially revealing when I asked the participants to say more about "hard". When a student described the test as hard, I received different meanings for hard across participants: hard as stress, hard as not knowing the answer, hard as having never taken the test, and hard as the length of the test. Easy had more to do with either liking the subject in which the student was being tested or in simply knowing the answers. Thus, hard and easy are not opposites.

Hard

One of the descriptions of hard had to do with the mental strain that stress put on students. It was interesting that the 5th graders, the older kids in this study, described hard as stressful. Sadie explained,

> *It was, kind of hard, cause you're stressed out because you know this stuff so you're making it hard (Sadie)*. She went on to repeat this same sentiment again later in the interview, *so then I just made myself think it was hard so most of stuff, so then it kind of was hard.*

Her perception was the test was not actually hard but her stress response to an important test was making it hard.

Sadie understood that she shouldn't be so stressed out, but she was. While at one part of the interview she felt she was *doing this to herself,* she mentioned the high stakes of the test adding to her stress as well, *I mean it is a big deal for grades and stuff like that.*

Sadie did not just experience her stress cognitively but also physically:

> *it wasn't like sleepy tiring it was like, you just, you're so stressed out about it because there's nothing really to be stressed out about you just get so tired from being stressed out you just want to lay down and go to sleep.*

Mark did not reference external reasons for his stress, but pinpointed his own inner drive to be the reason behind his stress. *Well, sometimes I'm nervous, I don't want to miss that many.* Mark also had a performance goal set for the TCAP Achievement Test; he wanted to score in the advanced level range. Mark's words reveals he understood the consequences of his stress, *like when I get nervous, I forget things*.

Kevin was the most explicit in how his stress made him feel. He was the only participant in the study to use metaphor to explain his nervousness:

> *when you're nervous, I just feel like um, I just feel like you are at a football game and you want to win really, really bad and it's like you're nervous when the roller coaster is really really high and you're nervous because it's really high and you are afraid of heights and when you are going down a roller coaster it kind of drops your stomach and that's what it feels like.*

Kevin's stress was felt, not just as an emotion, but, like Sadie, was felt physically, in his stomach. Kevin did describe this very intense stress subside however, midway through the test,

> *when you are halfway through the test it's when you start to get used to it and your aren't nervous anymore and um, when you are not nervous it feels good, cause it's like you are the smartest kid in the world.*

Hard was also described as not knowing the answer to a question on the test or not understanding the question. Carrie described this phenomenon most explicitly,

> *Some of the questions were kind of hard and some of them were kind of easy. It would depend on whether or not I knew it really well or didn't know it very well.*

Students cited confusion caused by not understanding the question as a reason why they couldn't answer or didn't know the correct answer, *some of them was like, kind of confusing.* When I asked Deshawn to say more about confusing, he continued, *it was like, it just didn't make that much sense to me, it didn't go through my mind right or something.* Carrie also used the word "confusing" to describe her experience, *like, I knew what it was but I just didn't know that was the word for it. So I was kind of confused like that.* Carrie explained her experience with a word problem in which one word was confusing for her. She also explained that not being able to get help from the teacher prevented her from answering the question.

Olivia also found questions confusing, especially on the social studies section of the test, *And social studies was o.k., but some things I didn't really get on the social studies.*

Olivia saw hard as experiencing the test for the first time.

> *And it like, when I first took it, I thought it would be kind of hard because, but…well, I thought it was 'cause, like I never took a – that was my first year, last year and I didn't know what it would be like to take it.*

Therefore, the anticipation of the test and taking it for the first time stood out to her. Kathy, Carrie, Sadie, and Frank all mentioned this phenomenon of taking the test for the first time and it making them nervous.

Some of the students described hard to mean the length of the test. *Well, hard is like, I mean so many pages and stuff (Mark).* The length of the test was elusive to some. One of the participants described taking the test for two weeks, when in fact, the test is taken over four consecutive mornings. Kevin also talked about the test taking longer than it actually took, *it took up the whole day and it was really long.* Actually, the TCAP Achievement Tests are only conducted in the morning, not all day. Student testing time for each day ranges from 136 minutes to 104 minutes, depending on the section being tested. However, whether they could remember the actual time they took it is not important, but the meaning they attached to it is. It was perceived as a long test, *when it's all together, all in one day, it's really long (Carrie).*

Students also talked about some negative physical aspects when describing the length of the test: *It's so long, your hand starts hurting (Carrie).*

Four of the students experienced the test as *painful*, *Yeah, and your hands get cramped and they're hurtin' a lot. Cause there's over 77 pages. The TCAP book is really big (Kevin).* Sitting in the chair for a length of period was also physically draining on one student. *You get kind of tired of sitting in the chair a lot (Chris).* As adults, I think we forget that sitting in a school chair for a length of time with no ability to get up periodically can hurt. Students sit in hard desk chairs with no cushioning. Each section of the test is timed. Student testing time for each day ranged from 136 minutes to 104 minutes, depending on the subject and section being tested. There are 8 sections of the test total, testing four content areas.

Carrie was concerned for the younger students who would have to experience the test in the years to come:

> *Well, little kids from first grade, second grade, and stuff, who are not used to, well, they're used to sitting down and being quiet and stuff, but they are not used to sitting down, being quiet, and being in their work for that long. Because first graders, they work and took a second grade test the next year and it would kind of been hard for them if they hadn't really learned how to sit still and be quiet for awhile.*

Carried continues the conversation talking about her own experience struggling with being quiet and staying still.

> *..even sometimes I get kind of jumpy because I've been sitting still for so long. And we haven't been able to kind of move around, but sometimes we get to stand up and stretch for a minute, but we have to turn our papers over and stuff. It's kind of weird because you sit down for so long you kind of can't move for long, but you stretch out because your muscles will get tired and your legs would probably get tired so you kind of have to stretch them out and stuff, so you don't stay like that for the whole day!*

Carrie also mentioned her eyes were bothering her during the test. *Your eyes would kind of get tired too, so you would have to get out of small distance, you would kind of have to look out so your eyes wouldn't be so tired*. She goes on to say

> "*The first day of testing that I did that, I asked if my teacher and my mom and dad, what they wanted me to do because my eyes would start hurting. They said just kind of look out the window for a minute because you might run out of time on the test.*"

Chris expressed his frustration in the length because he felt the pressure of the time limitations.

> *And the thing where you had the where you had a certain time to do it and it was kind of frustrating.*

When I asked him to say more about frustrating, he continued,

> *You have so little time to do it and if the time goes out then you can't go back and do that so you have to kind of guess on some of them and some of them you just know.*

Chris's frustration was not having enough time to go back to check on the ones he wasn't sure about or didn't feel he had enough time to figure out before feeling the need to move on due to the time pressures of the test.

Easy

Students described parts of the test as easy because they knew the answers to the questions.

Well, it was like, it wasn't really challenging because it was already what we learned.

Frank thought the Math portion was easy,

I missed one question, I think, of the, about that expanded form, how you got to write the number in, expanded form, and that's all. This made the effort expended on the test much less, *I didn't really have any trouble with it (Deshawn).*

Some of the students talked about the fact that liking the subject or task that they were being asked to do during the TCAP Achievement Test made it easy. *Well, I think the easiest was math and the reading because those are my favorite subjects (Olivia).* The fact that liking a subject made it easy has a lot to do with an increased self-efficacy of their ability to answer questions in that subject. Students described this self-efficacy in how much they enjoyed the subject of the questions. *Well, the social studies, it's fun, because I like social studies, answering questions about social studies. That's my favorite subject so it's always fun for me (Sadie)*. Kevin echoed this idea when he talked about his enjoyment of solving the math problems,

> *It was fun because in Math, well, I like doing math, subtracting and adding, and social studies, it was fun to learn about new stuff. And in science, I like to learn about plants and I like, um, to do projects and writing, I like to write about books and I like to read books.*

Theme II

Even though most of the students interviewed were only in third grade when they took the test, all talked about using strategies during the test.

Yeah, you have to really try hard (Kevin)

The most prevalent strategy across the interviews was guessing. *Sometimes I had to just guess 'cause it was just too hard (Kevin)*. Students are told to guess the answers if

you do not know it as you are not penalized for guessing (Tennessee Department of Education, 2010). According to Kathy, one teacher wanted her to choose a specific letter to apparently increase her chances of getting it right,

> *Because I didn't know the answer, and the teacher said last year, like, if you don't know the answer, just like, pick C, because that's what they should put the answer on. And sometimes I just guessed and sometimes I just put C."*

Students also used other strategies besides guessing, *I skipped it and went on to another one and I then checked over in case I missed one and made a mistake*. Chris talks about skipping over the ones he doesn't know which can both allow students another chance at the question later. It is also an important strategy with regards to time management. Instead of spending a lot of time on a question a student doesn't seem to know, then skipping over it can allow them more time on other questions. The student also talks about checking over his work which is really important especially in the math section of the test, when simple mistakes can keep one from answering questions correctly.

Mark described the process he went through in answering a question, *Well sometimes, like when I get nervous, I'll forget some things. I try to think back about it and how the teacher taught it to us*. He describes this strategy more explicitly in the next quote. *I just, like, take a minute to think about it, then I'll remember it. I try to have very much self-confidence as I can (Mark).* While Mark was nervous, he was able to overcome this by using a successful strategy to help him perform on the test. First, he took the time to think about the answer until it came to them In this instance, thinking about it wasn't enough. The student knew that to get the question right, he couldn't react in a negative

way to a question that he didn't know right away. If he had self-confidence, the answer could come to him if he thought about it.

Kevin talked about trying really hard as a strategy to do well. *I still tried my hardest, give 110% and you always have to try hard and keep trying and trying 'til you get it (Kevin).* If trying hard still didn't yield results, he used another strategy, "*yeah, you have to really try hard and um try hard to figure out the problem. If you don't know it, the problem, you can go ahead and come back to it*." Unlike Mark, Kevin didn't seem to need self-confidence in order to continue working hard, but focus was a key component to his success. "*…all you have do really do is concentrate*."

Theme III

It decides…well, I don't know, but it seems… (Carrie)

Students - while they knew the test was different, that they needed to try hard, and that their teacher had been preparing them for the test for several weeks prior to taking it - were confused about why they were taking the test. Several ideas were proposed by the students as to the purpose of the test. Carrie suspected that it affected you're grade in school, *It decides…well I don't really know, but it seems like it would count for way more of your grade than first graders would be because we know more so they expect us to know more because we are older. Unless you had a problem or something like a mental problem. But they expect you to know bigger words like my teachers at school do. So they want that bigger more amount of your grade (Carrie).*

Carrie's sister, Sadie, agreed and explicitly cited her teacher as the source of her information,

*Well, my teacher said it's like some part of your grade and it helps you with when you get older and stuff like that (Sadie).*She also restated later on in the interview that the test was a bigger part of your grade when she was explaining why she was so nervous going into the test, ***I*** *mean it is a big deal for grades and stuff like that (Sadie)*.

Frank was sure that the tests helped teachers understand what students know and didn't know so they could help him improve on skills that he did not know, *So they will know how smart they are and what they need to learn, like if I didn't know what 6 X 12 was or something, then I need to practice my multiplication.*

Mark was able to explain the different levels of the test. Mark was the only one who talked about the results of the test in technical terms. *Well, there's usually an advanced level, and if you aren't just right for it, you're below level, and I try to get advanced level…it's when you like you get more than what you are supposed to, more than what you have to get.* In each section of the test, there are different categories. For example, within the Reading/Language Art Test, there are seven categories that are scored: Content, Meaning, Vocabulary, Writing/Organization, Writing Process, Grammar/Conventions, Techniques and Skills. In each of these categories, students can score in three different levels: Advanced, Proficient or Below Proficient. The Advanced Level is the highest level that you can obtain.

One student said she did not personally see her scores, but knew that her whole class did well because they were given a pizza party as a reward, *They never get us our test scores from our TCAPs but I know our test scores were really hard cause our*

class got the bestest scores in third grade (Olivia). The Tennessee Department of Education in the past, has sent two copies to the school after the tests are scored, according to the Supervisor of Research and Evaluation for Knox County Schools. One was for the student's cumulative record and the other was to be sent home (J. Beckett, personal communication, July 19, 2010).

Kathy believed that the TCAP Achievement Test was something that could keep her from moving on to the next grade: *I felt like I was going to pass it and move on to the fourth grade and I did (Kathy).*

John Beckett, a Supervisor of Research and Evaluation for Knox County Schools says that test scores are used primarily for accountability purposes school wide. Schools may use the tests as additional evidence that a child may need to be held back a grade, but this is a very rare occurrence (J. Beckett, personal communication, July 19, 2010).

Summary of Findings

Students were all grounded in their experience of taking other tests besides the TCAP Achievement Test and all of the themes were based on this experience. The students knew that the test was more important/weird than other tests. All described different aspects of the test that were important/weird: the differences in teachers, the quiet, and getting a snack were all listed as things that stood out to the participants.

Students also described the test in terms of how hard or easy it was and gave meanings to what hard and easy meant to them as test takers. All the students shared their strategies they employed when taking the test. Finally, students did not understand

the purpose of the test. However, this did not seem to anger them or create any hard feelings about their overall experience of the test.

Chapter V

Summary and Conclusions

In a prior phenomenological study that explored teachers experience in high stakes testing environment, dire concerns were expressed about effects of high stakes testing on students. Teachers were so concerned they made statements like, “it kills you” to describe their emotional response to watching students in a high stakes testing environment. They were explicit in their conclusions about high stakes accountability, “What’s sad is that it doesn’t work for the kids.” Teachers used strong negative words to describe the test such as “horrible” and “demeaning” (Crisp, Williams, and Greenberg, 2006).

The purpose of this dissertation was to understand the lived experience of students taking a high stakes test. The phenomenological method developed by Dr. Howard Pollio (Thomas and Pollio, 2002, Pollio, Henley, and Thompson, 1997) at the University of Tennessee was used to gain a deeper awareness of what it’s like for an elementary school student to take a high stakes test. Nine elementary school students from an after school program at a recreation center in a city in the Southeast voluntarily participated in the study. Participants ranged in age from 9-11 and were either in the 4^{th} or 5^{th} grade.

The following sections of Chapter V provide a summary of the findings from the study, a discussion of the findings, implications for practice, recommendations for future research, and conclusions.

Summary of Findings

Despite teachers concerns that students are suffering under the current accountability system, findings from this study, while not refuting teachers' thoughts and feelings, did not find that students harbored such negative feelings about the high stakes tests. In general, there was an ambivalence that existed in the students' experience to the test. The test, they knew, was more important than other tests and students were able to talk about the differences of the TCAP Achievement Test week, some good, some uncomfortable, some just weird. While the students explained that the test could be hard, they also thought it was easy, and they used strategies to help them take the test.

Students experienced good things about the test, like doing well on the subjects that they liked, and getting to have snack during the break. They were also aware of their teachers who were not physically present to ensure teachers wouldn't cheat. Even though the test administration procedures stipulated that teachers were to read a statement to students about the purpose of the test, the students were still unclear on the exact purpose of the test. But, students were clear that the test was important, and had definite theories about why it was important. In their own words, students depicted a culture of high stakes tests exams.

A Culture of High Stakes Exams

The characteristics of a high stakes exam culture as experienced by the students are represented in the Ground and three themes (Ground) aspects of this test that made it different than other tests: *It was just like another test, but like more important/ weird*; (Theme I) reactions to the actual test: *It was kind of hard and kind of easy*; (Theme II) strategies that help you on the test: *Yeah, you have to really try hard*; and (Theme III):

The purpose of the test: *It decides, well, I don't know, but it seems….* The physical aspect of the test, while it wasn't a theme, was also a part of the test experience for some of the students: *It's so long, your hand starts hurting*.

Students described the environment in which they took the tests that made it a more important/weird test: They talked about their teachers having to switch classrooms, their teachers under suspicion of cheating, and their teachers not being able to help them. They mentioned the extra quiet during the test, and its significance. For one student, quiet was an action; having to *be quiet*. For another, it was a state of how the test was: *peace and quiet*. Finally, juice and snacks were part of the culture of this test. During breaks they were given a special snack and candy. All of the students felt that getting snacks during the break was a good thing; a reward for all their hard work: *That was a good thing we got to look forward when we finished the test...*

The participants talked about their experiences as they took the test, which included physical and cognitive reactions to the test as being either hard or easy. The students' reaction to the questions depended on whether they knew it or not. If they knew the problems or liked the subject they were being tested on, the test could sometimes be "easy" and was even fun. While experiencing the test as "hard" could be describing the fact that the student didn't know the answer, the adjective was also used to describe the length of the test which was physically uncomfortable, or being confused when reading a question. During the test, students talked about the strategies they employed, including guessing, trying hard, and taking time out to think about the question.

All of this produced or influenced the meaning of the test. The importance of the test influenced the strategies they used. The test had several purposes and

consequences, according to these students. Three of the students thought it counted toward your actual grade. Two of the students thought it helped teachers understand what you needed to learn. One participant thought it decided how smart you are, and another student thought it helped when you got older in some way. Finally, one participant thought it decided whether or not you would go onto the next grade.

Discussion

The following discussion presents the current study's findings in relation to the literature on effects of high stakes testing as well as related literature.

The Ground of the Experience: **It was like, just another test…**

The ground of the experience is that which all of the other themes 'stand out of'. In other words, without the ground, the other themes would not exist (Thomas & Pollio, 2002). For students taking a high stakes test, their past experience of taking other tests during the year helped them to understand the other aspects of the test that stood out to them about this *important/weird* test.

A regular test during the school year is usually given by the classroom teacher, and the teacher is able to assist students in understanding the question without giving them the answer. No snack is given and while there is a modicum of quiet, it's not *so quiet.* The regular classroom test is not *such a long test* that it makes your hand hurt. The ground in this study assists the understanding that what stands out to them is different, but it also is revealing in terms of what they did not talk about. While they told us what stood out to them, what did not stand out to them were the very elaborate and strict guidelines of the test.

In a standardized test, students are given their tests, told to fill out the identification information of the test, and also given very careful and exact directions that include when to open up the booklets, what page to turn to, when to read the instructions, where and how to mark your answers, when to begin the test, and when to stop. Students are not allowed to leave the room accept during an assigned break. They are not allowed to talk to fellow test takers and must stay in their seat until the assigned break. Teachers are bound by a strict list of guidelines that must be followed exactly, and even have to sign a contract in which they risk losing their jobs among other consequences if they do not (see Ch 4). The test also requires students to fill in bubble answer sheets to allow for electronic scoring. One bubble, or small circle, must be filled in completely with no stray marks. Test directions warn the students that incorrectly filling the bubble sheets out could alter their score.

According to students, the specific rules and instructions did not stand out to them, even though these rules and instructions are different from other classroom tests given at any other time of the year. What was meaningful to them were other aspects of the test environment that were different than other tests: change in teachers; teachers potentially cheating; the quiet; and the snack as well as the content of the test: whether it was hard or easy; strategies they used; and why it was important. Students were reacting to a cultural mismatch. Cultural mismatch in an educational context is a term usually used to describe a difference in culture between home and school (Harris, 1991). However, in this case, the cultural mismatch is between ‘regular’ testing – testing that occurs throughout the school year – and TCAP Achievement testing. Cultural mismatch can interfere with

academic achievement (Garcia, 1995). It can also lead to negative feelings toward school, like anger and alienation (Ormrod, 2008).

A plethora of position papers and studies have articulated the changes that high stake accountability systems have had on schools in the United States (Amrein and Berliner, 2002, Nichols & Berliner, 2008, Bickham, Burns, and Monahan, 2001, Gordon and Reese, 1997). These studies cited negative consequences to the school culture that had an impact on student learning including the narrowing of the curriculum, an increase in tests during the school year, an increase in teaching test taking strategies, decreased motivation in students, and a change in role of the teachers.

One of the most revealing and thorough studies conducted by Valli et. al. (2008) analyzed the impact of high stakes accountability testing in three elementary schools. In the two year study, they observed the effects of high stakes testing, and found that the schools created their own test-taking culture, where test performance became the ultimate goal of instruction.

> In test taking cultures, learning is supplanted rather than supported by assessments. Schools participate in gaming strategies to avoid adverse consequences, and teachers reshape instructional activities to mirror standardized tests. As a result students often learn less than when learning, not testing, is the explicit goal (p. 25).

This study, instead of looking at the overarching system, was focused on understanding the experience of students taking a high stakes test. I did not ask students about their daily classroom environment. Students answered my question by focusing on the actual week of the test and not the culture of a high stakes testing environment

throughout the school year. Evidence from the interviews does confirm some of the positive and negative reactions to high stakes testing found in previous studies

Students talked about the differences in their teachers. Students talked about their teachers doing two things as they took the test: 1) not being there, but wanting to be there, 2) explaining that they couldn't help with the test. This change in role of the teachers is captured best in Kathy's comment, *the teachers, they couldn't help us.*

Linda Darling-Hammond labels our current culture of accountability as a 'Bureaucratic Approach' to education. "The fundamental assumption is that this process, if efficiently administered, will produce the desire outcomes" (p. 47). Efficient administration is played out in an ever increasing role of administrative controls in the classroom that affects curriculums, classroom schedules, and testing regimens (Darling-Hammond, 2010).

Standardized tests are not only a part of the bureaucratic approach but really epitomize the bureaucratic philosophy. Standardized tests come with a manual and a prescriptive checklist for teachers to strictly follow. Procedures must be closely followed as denotes the very definition of a standardized test.

As an educator, the teacher's responsibility throughout the regular school year is to assist students in order to help them achieve. During a regular classroom test, if a student doesn't understand a question, the teacher would help a student to understand it without giving away the answer. In fact, good practice encourages teachers to create an environment where students feel comfortable asking questions if something is not clear (Ormrod, 2008). In a standardized test, however, this important teacher role is eliminated

in order to maintain the integrity of the test. The teacher's role changes during a standardized test, from being helpful to being unhelpful.

As participants explained in this study, the teachers' role also changed from trusted teacher, to a teacher under suspicion of cheating. While students do not think their teacher would cheat, *others might.* Teachers not only are under suspicion, but must switch classrooms or be monitored by another adult.

While the area where this study was conducted has not experienced a cheating scandal in several years, one study has shown that certain reasons would justify cheating. A majority of teachers surveyed believed that if students benefitted or the administration of the test was felt to be inappropriate, teachers would consider helping students in ways considered cheating, such as helpful hints, rewording items, and teaching to the exact test (Wellhousen & Martin, 1995).

The possibility of teacher's cheating has increased under a high stakes environment. How do children process this information? It was clear in this study that the students were protective of their teachers even though they were very aware of the reason their teachers were not in the room. It seems that children look to 'other teachers' as the cheaters to help them process the fact that people may think their teachers could cheat: *because um, otherwise your teacher might, but probably won't, (Carrie).*

Unfortunately, administrators who rely on these test scores for accountability purposes may be curing one problem but causing another by taking out the students' teacher. Research done in psychological testing shows that students' relationship to their proctor affects their scores. Children actually perform better when the teacher is known to them (DeRosa & Patalano, 1991). A meta-analysis conducted found that test scores

increase by .28 standard deviation if the person taking the test knows the examiner (Fuchs & Fuchs, 1986).

This standardization and more rigid/externally determined set of rules may also increase the next negative consequence of high stakes testing: an increase in test anxiety, revealed in theme I.

Theme I: **It was kind of hard and kind of easy.**

Test anxiety is another detrimental outcome of high stakes standardized testing (Fleege, et. al., 1992). Teachers report that students experience an increase in test anxiety when tests are tied to high stakes (Gordon & Reese, 1997). Test related stress includes two components: 1) the worry that comes in anticipation of a test and 2) the emotions that are triggered during the taking of the test (Jones, et.al., 1999). Students in this study clearly expressed both components of test anxiety. Students also exhibited physical responses to stress such as tiredness and a nervous stomach.

This study confirms prior research that also found students suffer from test anxiety while taking a high stakes test (Wheelock, Bebell, and Haney, 2000, Triplett & Barksdale, 2003). It also confirms teacher's reports that students experience an increase in anxiety during these tests (Gordon & Reese, 1997, Jones et. al., 1999, Lattimore, 2005).

Test anxiety is a serious issue with some students and can lead to lowered achievement, decreased social functioning, and lower feelings of self worth. Test anxiety also has an increasingly negative accumulative effect. Once a student associates himself or herself as being 'a bad test taker' student's anxiety is likely to increase (Jones, et.al., 1999).

The fact that the time and length of the test caused physical pain and discomfort to the students was a compelling finding in this study. Students' hands hurting could also be attributed to graphomotor issues. Students with graphomotor disorders require increased cognitive and motor effort in order to control their pencil during the test compared to students without the disorder and this would certainly affect their performance (Waber & Bernstein, 1994). Intervention in reading has increased as a result of NCLB; however, students with poor writing skills do not receive the same scrutiny. Most children with graphomotor disorders go undiagnosed and some estimates say that up to 25% of children could have the disorder (Bates, 2010). Others may not meet the criterion for a graphomotor disorder, but may grip the pencil in such a way that the hand quickly tires when writing. Teachers should be able to identify those students during the year who hold their pencils at odd angles, who have poor handwriting, or avoid writing. When students say they dislike writing, this is also an indicator that there is an issue (Waber & Bernstein, 1994).

On the positive side, evidence of self-efficacy was also found in the interviews. Self-efficacy can have a positive impact on the performance of students taking tests. Self-efficacy relates to psychological concept that how competent you feel about something will determine how well you like something and the effort you expend on it (Bandura, 1977). Self-efficacy is highly susceptible to environmental issues. They are: performance accomplishment or failure; vicarious experience; verbal persuasion; and physiological states (Bandura, 1989).

Evidence in both the Wheelock, Babell, and Haney (2000) study also contained evidence that some of the student's drawings depicted self-efficacy. Studies have

focused on self-efficacy pertaining to academic motivation. Two measures of self-efficacy in particular have been researched: rate of performance and expenditure of energy. Research found that perceived self-efficacy is positively correlated with math performance (Schunk, Hanson, and Cox, 1987). In addition, self-efficacy and self-rated mental effort was positively related to learning difficult text (Salomon, 1984). Studies have also found that instruction can have strong effects with regards to self-efficacy (Bandura & Schunk, 1981).

Self-efficacy and anxiety are both affective or emotional factors that can attribute to performance on high stakes tests. Triplett and Barksdale (2005) concluded that students expressed negative emotions when asked to draw and describe. “The recurring words that expressed emotions included: ‘nervous,’ ‘mad,’ ‘sad,’ ‘frustrated,’ ‘hate,’ ‘confused,’ ‘bored,’ ‘tired,’ ‘sweating,’ and ‘sleepy,’” Out of all of these emotions ‘nervous’ was the one reported most often (Triplett and Barksdale, 2005, p.244).

In this study, all of these emotions were found during data analysis with the exception of the words mad, sad, hate, and sweating. However, what stands out in my study is that students connected each of the remaining emotions they did describe to certain phenomenon of the tests. For example, those that used the word nervous in their interviews connected the nervous experience to the idea that it was their first experience with taking this type of test. It was a feeling that students felt at the beginning of the test, but not in the middle or the end of the test.

The experience of frustration was used to explain their reaction to the limited amount of time that students had on the test. Finally, confused was explained as what they experienced when they didn’t understand the question or didn’t know the answer.

Some students used the guessing strategy when confused.

Theme II: Yeah, you have to really try hard

Teachers typically instruct their students on strategy use before the test to help students perform their best. In fact, one of the criticisms of high stakes tests is the time used for test taking strategies and the increased assessments during the year to prepare for the end of the year tests take away valuable class time that could be spent learning new knowledge (Linn, 2000).

Some test strategies that are taught to students can promote learning how to learn and higher order thinking skills. Others can reinforce less desirable learning habits (Paris, 1991). Both types of strategies were revealed by the participants in this study. Kathy explained, *and the teacher said last year, like, if you don't know the answer, just like, pick C..."*

Kathy is explaining the guessing strategy, a strategy that has negative consequences to learning. However, she also talks about an effective testing strategy that can have positive ramifications to learning: *like on the reading, some I got if I go look at the story.*

Other students talked about thinking about the answer until it came to them. *I just, like, take a minute to think about it, then I'll remember it. I try to have very much self-confidence as I can* (Mark). In this instance thinking about it wasn't enough. The student knew that to get the question right, he couldn't react in a negative way to a question that he didn't know right away. If he had self-confidence, the answer could come to him if he thought about it.

The evidence in this study confirms findings in the Jones and Egley studies that teachers who are in higher stakes environments are more likely to teach test taking strategies (Jones & Egley, 2007). Pedulla (2000) found that 92 % of teachers in schools with high stakes tests spent time on test preparation, and over 60% of teachers were reported using test-specific preparation materials given to them by the state. Other studies have found that test preparation, including teaching strategies, reduced the amount of time available for instruction (Jones, Jones, & Hargrove, 2003, Hoffman, Assif, & Paris, 2001, Jones & Egley, 2004). In Texas, teachers reported spending between 8-10 hours a week year round on instruction for the end of the year standardized test (Hoffman, Assif, & Paris, 2001). Gordon & Reese found the same scenario in their study and also found that the hours devoted to the Texas assessment increased significantly up to 3 months before the test. One teacher reports “Our curriculum was totally different a month before TAAS; it didn’t consist of much social studies and science at all. Writing was dropped because the writing test was taken in March” (Gordon & Reese, 1997, p. 355).

Studies focused on students’ perceptions did not find evidence of strategy use in their studies. However, one study on high school students focused on student reactions to test preparation for a high stakes test. The student’s interviewed expressed frustration about the test preparation in that it reviewed the information too fast. The researcher observed that the teacher instruction changed from a focus on conceptual understanding to a focus on specific procedures and protocols (Lattimore, 2005). Students in this study employed positive strategies, like thinking through a problem and trying to maintain confidence throughout the test. While guessing is not typically a good test taking strategy,

in a test like the TCAP Achievement Test, guessing does not count against you. Therefore, in this context, it was a good thing to at least attempt an answer.

While evidence from the interviews found students were well versed in effective strategies for taking the TCAP Achievement Test, the students were not as clear on the purpose of the test.

Theme III: It decides…well I don't know but it seems…

With the variety of purposes given by students, it is clear the purpose of the test was very unclear. Students had a variety of ideas about the test. In the Teachers Administration Manual for the TCAP Achievement Test (2010), the teachers are advised to give the following statement on the purpose of the test: "The purpose of taking an achievement test is to find out which skills have been learned and which skills need further development" (p.8). Most of the manual, however, is dedicated to a script and a list of standardized procedures to follow. Laurie Driver, the Supervisor of Testing for Knox County Schools trains the test administrators of individual schools. The training for administrators is focused on ensuring the security of the test due to the dire consequences that can befall a teacher if he/she intentionally or unintentionally varies from the test instructions (L. Driver, personal communication, July, 23, 2010).

This lack of clarity on the purpose of the test may also be due to the lack of overall training for teachers in schools with regards to measurement issues and testing (Stiggins, 1994). Literature has been published as a response to this increase in testing in an effort to diminish possible negative effects of these tests. The Association for Assessment in Counseling has published a document called the Responsibilities of Users of Standardized Tests (RUST). The intent is to help those who implement standardized tests

follow “responsible testing practices” (Association for Assessment in Counseling, 2003, p. 385). RUST calls for test users to be highly qualified, with the technical knowledge, the need to understand the testing instruments, and results of that instrument. RUST also calls for the test administration to include information that outlines the purpose of the test.

This research study reflected and confirmed effects of high stakes testing that have already been discussed in other research and position papers. It also added to existing research. Based on the findings of this study, some suggestions for additional research and recommendations to improve practice are outlined in the next section.

Considerations for Practice

Based on the words of the participants, suggestions for research and practice can be made in order to improve student’s experience that involve a human touch to testing, as well as increasing awareness and the identification of graphomotor difficulties that can be problematic during *such a long test* (Bates, 2010).

That tests must be implemented using a standard administration has been a long held belief in both psychological and educational testing. However, standard doesn’t have to be equal. Nor can we say that just because something is equal means that something is fair. Levin (2009) purports that for reform to be effective it must include the following: “changing teaching and learning (and assessment) practices, improving teachers’ skills to do so, strengthening leadership capacity, improving student engagement in learning, and reaching out to parents to support their children’s learning” (p. 262).

While standardized proctor procedures are put in place to make the test more valid, the rigidity of these standard procedures created a cultural mismatch that may impact the students’ performance. Particularly, the relationship between the teacher and

student is altered as they are not in the room with the students. Even if they are in the room, they are reading mechanically from a required list of instructions and cannot assist the students when students have questions.

When student are experiencing cultural mismatch, Igoa (1995) makes the recommendation to maintain continuity between your room and the “home-room” that could be used to aid students experiencing cultural mismatch between what they experience during their informal, teacher-made tests and that of a TCAP Achievement Test. Teachers and proctors should acknowledge their own regular culture and explain how it will be different from the TCAP Achievement Test culture. However, while standardized tests require a standard procedure to follow, some standardized procedures, like not helping a student with a question, seems to be an unhelpful practice that has the potential to damage the relationship between student and teacher. In this same vein, the awareness that teachers would cheat is another potentially harmful issue that should be addressed.

The following are recommendations to helping students both adapt to this new culture as well as to ensure a culture of caring between children and their teachers:

- Provide a proctor to monitor test administration while allowing students’ own teachers in the room in order to give students the best possible environment to do well on the test (DeRosa & Patalano, 1991, Fuchs & Fuchs, 1986).
- Train teachers on test anxiety reduction methods such as breathing techniques and positive self-talk (Ergene, 2003).
- Help students understand the purpose of the test. Facilitate a discussion about the purpose of the test prior to the test day where kids list all of their

assumptions and beliefs of why they are taking the test as well as any consequences. Then give them the real purpose of the test and help them to compare the purpose to their answers. Group discussion of a topic that may have several different meanings allows students to gain a better understanding of a topic by giving them an opportunity to evaluate their own assumptions against others' (Ormrod, 2008).

- The idea that teachers may cheat is detrimental to both the school culture and the child's perception of the teacher. Students should not be made aware that the main purpose for security measures is to keep a teacher from cheating.
- Students should be told why they are being read to by script in an effort to help students process the change in their teacher (Igoa, 1995).
- Teachers should be given a standardized method to assist students when they don't understand a question on the test so that students feel that they are getting help.
- Students with writing issues should not have to fill in the bubbles in the test but should be allowed to simply circle the answers, as students with IEPs or 504s are already allowed to do.

Recommendations for Future Research

This study revealed a high stakes testing culture that has the potential to negatively impact a student's ability to do well and to increase psychological distress of students. More research needs to be done to explore the cultural mismatch of schools under high stakes accountability. Culture has a very powerful impact on our minds and "provides us with a toolkit by which we construct not only our worlds but our very

conceptions of ourselves and our powers" (Bruner,1996, p.x). It is through interacting with others that children discover the ways and norms of the culture in which they live. During standardized tests, we suspend this interaction between teacher and student, limiting or substantially changing it to meet the needs of validity and test security. It is questionable, then, to ignore psychology and good educational practice all in the name of standardization. If we want students to feel that school is a respectable, fair, caring, and meaningful place to learn, it seems some of our practices during standardized tests fall short.

Additionally, while accountability reform is currently an entrenched part of our educational system, Linda-Darling Hammond (2009) concludes from her years of researching bureaucratic reform that we must look at the power differential of the high stakes testing system:

> Until authority for making decisions is granted to those who have responsibility for performing the work – those who are living with these decisions on a daily basis – reform of practice cannot occur. But ensuring that good decisions are made depends, once again on the availability of solid professional knowledge and strong moral commitments within the school (p.55).

- Thus good professional practice can occur when well trained teachers not only have the responsibility, but also the power to carry out sound educational practices. The following bullet points are suggestions for future research: Learning from the areas of psychological testing, general educational psychology principles, and other protocols should be explored in an order to improve the humanity of high stakes testing. General teacher education includes the exposure to educational

and psychological theories on student affect and how it can improve or become a barrier to learning. However, the standardized tests' rigid and controlled methods, the length, and the time limits do not seem to take students' emotional needs into account. In psychological testing, observations of the student's behavior, effort expended, and any other pertinent observation made of the student during testing are always included in the educational report. These observations are used as evidence in order to conclude that test results are a valid indicator of one's ability. Students' perceptions of assessment and their motivation must be considered when looking at and making judgments based on test scores (Paris, 1998).

- It would be helpful to explore how teachers address student distress in a high stakes testing situation compared to a regular test. The focus on test security is important as it leads to grave consequences for teachers if not followed. However, student distress is also important and has the potential to not just effect scores of students, but to break from the norm in how teachers care for children. The Teacher Administration Guidelines (2010) distributed by Tennessee Department of Education is intent on securing the test, but has no advice on how to address a student dealing with distress. In other professions that utilize standardized tests, such as counseling and school psychology, professionals are guided by a code of ethics that includes a section on client welfare.
- On a related issue, it would be interesting to explore teachers understanding of standardized tests, particularly on test-taking considerations, test-administration procedures, and test wiseness. Most teacher education programs do not require training in assessment (Stiggins, 1994). For example, the University of Tennessee

program requires students seeking a teaching certificate to take a class on research methods and statistics, but not on assessment ("CFS Core Course Requirements", 2010). However, they do get exposed to test measurement -for one week -in a required 3 hour course: Educational Psychology 401.

- The insight gained from this study suggests that more research is needed using this approach with larger groups of students in more schools and districts and states at more ages as well as paying special attention to those most at risk including low-socioeconomic groups, minorities, and special education students.

Personal Reflection on Interviewing Children

Interviewing students was an enlightening experience to my practice as a researcher. I was surprised at first at how simply my participants talked about their experiences. I also noticed a lack of judgment that I haven't experienced in adult interviews. This stood out to me because the test was such an artificial construct in which they had no power. I assumed they would have much stronger feelings about it.

Certainly, this simplicity and lack of judgment could have been influenced by the conditions of our interview. Sitting down with an acquaintance in the director's office of a recreation center and being asked an open-ended question would be a unique experience for these kids. In addition, the students may have found more to say if they had just taken the test (six months had past between the test and the interviews). These factors may have been the reason for the short interviews.

However, I did take precautions in order to glean as much data as I could from this study. A previous pilot study taught me invaluable lessons about interviewing children. In the last study, I interviewed the students in a school, and none of the students, prior to my

interview, had ever seen me before. Therefore, I changed the place of the interview to what I felt to be more neutral territory- a recreation center. It was also a place where students were at least acquainted with me. I also learned from this previous pilot study that I needed to do a better job and communicate my intentions, so that they had a better understanding of the purpose of the interview. Previously, I noticed students in the pilot study were trying to give me the right answer to my question even though there was none. Therefore, I added a short talk with the students before we began the interview that I was only interested in what they thought. That I had interviewed teachers and found out what they thought about the test, and now I was *really* interested in what they could tell me about taking the TCAP. Achievement Test I also told them that I was a student, just like them, and was interviewing them as part of a school project.

Certainly, my care in creating a more comfortable, more purposeful interview would not have eliminated all of the reticence that a child may have talking with an adult. However, I was still surprised during the interviews the participants didn't seem to have strong feelings towards a phenomenon that seemed, to me, so restrictive and that had been constructed for them. High stakes testing is an adult construct. They did not help create nor do they have a voice in the development, the implementation, or the consequences, yet they are at the center of this political, hierarchical, top down construct that affects an increasing amount of their school experience.

My understanding of their simple, ambivalent responses to my question grew during the research group analysis of the students. It was also later confirmed as I went over my findings with Dr. Pollio. In both of these discussions, as we pored over the transcripts, we all saw a sense of confusion that might help explain their lack of detail or

feeling in their answers. This confusion is exemplified in the Ground: *It was just another test* and Theme I: *but more important/weird.*

As a student and researcher in Educational Psychology, I have learned and experienced a very simple yet profound and often ignored idea in education: that for students to see something as important, it must have meaning for them. As elementary students taking the TCAP Achievement Test, while the test was not meaningless - as they all understood it had some purpose- it was confusing. I believe this confusion helps me to understand the later angry and resentful responses about high stakes tests we see in other studies done with older students.

Another thing I noticed that was not in the students' interviews is the lack of metaphor in the student's descriptions. In past studies, metaphor has been of vast help in kneading out the meaning of the experience. Unfortunately, out of the nine interviews, only one child, Kevin, used metaphor to describe his experience. However, this one child gave some extremely moving metaphors, and while these have already appeared in the research findings, I wanted to lay them out here for further emphasis.

Kevin used very powerful analogies to describe his experience of being nervous:

> *And like, when you're nervous, I just feel like um, I just feel like you are at a football game and you want to win really, really bad*

This metaphor really helped the research group understand what Kevin meant by nervous and shows the intensity of Kevin's nervousness. His nervous was the nervous you get when you are about to perform. It's also the type of nervous that you get when you can only get one chance at something. It also illuminated his strong motivation to do well on the test.

and it's like when you're nervous it's like you're nervous when the roller coaster is really, really high and you're nervous because it's really high and you are afraid of heights and when and when you are going down a roller coasters it kind of drops your stomach and that's what it feels like.

This metaphor helped me to understand that Kevin was physically feeling nervous, and again, the intensity of this feeling stood out to the research group.

Kevin went on to speak metaphorically about his experience at the middle of the test, *when you are half way through the test it's when you start to get used to it and you aren't nervous anymore and um, when you are not nervous it kind of feels good, cause it's like you are the smartest kid in the world*.

My experience of talking with elementary students has given me a deeper understanding of children's responses. While on the outset, and looked at individually, children's interviews may seem at first simple, lacking the detail that I, as an adult researcher, feels is necessary for a successful study. However, collectively, the interviews begin to develop 'legs' from which real discovery can be made. "Out of the mouths of babes" is an appropriate colloquialism to describe this discovery. This research experience confirmed an original hypothesis and past research: students are an important voice that should be included in educational reform.

Conclusions

While my question, "what was it like for you when you took the TCAP test" seemed on the outset narrow, as it asked for them to talk specifically about the four days they took the test; students, in their responses to this question, described an educational culture that confirmed existing research. This culture they described was grounded in their past experience of taking other tests. They were able to reveal the context in which they took

the tests that was important or stood out to them: their teachers having to switch classrooms, the quiet during the test, and how they were given a special snack and candy. They talked about the actual experiences of the test, their physical and cognitive reactions to the test as well as their perceptions of the test being either hard or easy. Something that was not revealed in prior research was student use of strategies. Student used strategies taught to them by their teachers to help them as they worked on the test. All of these aspects of the experience of the test produced or influenced the meaning of the test that has several purposes and consequences, according to these students.

A growing number of phenomenological studies in applied fields have also demonstrated the ability to illuminate valuable information about practice and its context (Thomas & Pollio, 2002; Greenberg & Orechkina, 2004). Like studies before it, this study on exploring the experience of students taking a high stakes test has shown that asking children about their experience can enlighten our understanding of standardized testing practices that impact test scores schools rely on so heavily to prove their efficacy as an institution. If we are to support our teachers and our schools in this era of high stakes, we should acknowledge the voices of our students who are taking them, and adjust our policy and practices accordingly.

REFERENCES

References

Adams, E. (2006). Scientifically based research in education: Epistemology and ethics. *Adult Education Quarterly, 56*(4), 239-266.

AERA (2000). *Online position statement concerning high-stakes testing in P-12 education*. Retrieved from: http://www.aera.net/about/policy/stakes.htm

Alvarez, L. and Corn, J (2008). Exchanging assessment for accountability: The implications of high-stakes reading assessments for English learners. *Language Arts, 85*(5), 354-365.

American Psychological Association (2001). Appropriate use of high-stakes testing in our nation's schools. Retrieved from: www.apa.org/pubinfo/testing.html

Amrein, A.L. & Berliner, D.C. (2002). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives, 10*(18). Retrieved from: http://epaa.asu.edu/epaa/v10n18/

Amrein, A.L., Berliner, D.C. (2002). *The impact of high-stakes tests on student academic performance: An analysis of NAEP results in states with high-stakes and ACT, SAT, and AP test results in states with high school graduation exams* (EPSL-0211-126-EPRU). Tempe, AZ: Arizona State University, Education Policy Studies Laboratory.

Amrein, A.L., Berliner, D.C. (2003). The effects of high-stakes testing on student motivation and learning. *Educational Leadership*, *1*, 32-38.

Amrein-Beardsley & Berliner, D.C. (2003). Re-analysis of NAEP math and reading scores

in states with and without high stakes tests: Response to Rosenshine. *Education Policy Analysis Archives, 11*(25). Retrieved from http://epaa.asu.edu/epaa/v11n25/.

Anderson, L. (2007). *Congress and the classroom: From the cold war to No Child Left Behind*. University Park, PA: Pennsylvania State University Press.

Association for Assessment in Counseling (2003), Responsibilities of users of standardized tests (RUST). In S.C. Whiston, *Principles and Application of Assessment in Counseling*, pp. 384-390. US: Thompson.

Association for Supervision and Curriculum Development (2005), Key Measures in the New Law, Retrieved from http://www.ascd.org/portal/site/ascd/menuitem.44cb9f9033aaf17cbfb3ffdb62108a0c/template.article?articleMgmtId=8a2c2fec6c462010VgnVCM1000003d01a8c0RCRD

Austin, T., Goals 2000: The Clinton Administration Education Program. Retrieved from: http://www.nd.edu/~rbarger/www7/goals200.html.

Axtmen,K. (January 11, 2005). When tests' cheaters are the teachers. The Christian Science Monitor. Retrieved from: http://www.csmonitor.com/2005/0111/p01s03-ussc.html

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215.

Bandura, A. (1989). Human agency in social cognitive-theory. *American Psychologist, 44*, 1175-1184.

Bandura, A. and Schunk, D. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social*

Psychology, 41(3), 586-598.

Barksdale-Ladd, M.A., & Thomas, K.F. (2000). What's at stake in high-stakes testing: Teachers and parents speak out. *Journal of Teacher Education, 51,* 384-397.

Bates, B. (2010) Writing wrongs: Treating graphomotor problems. Clinical Psychiatry News. Retrieved from: http://findarticles.com/p/articles/mi_hb4345/is_6_34/ai_n29281348/

Bickham, T., Burns, P., Monahan, D. (2001). *Emphasis placed on current assessment procedures affects the way teachers teach* (Thesis, Saint Xavier University, 2001).

Bierbauer, C. (1996). *Now that Johnny can read, can he get a job with IBM? Governors, business leaders tackle education woes*. CNN Interactive. Retrieved from: http://www.cnn.com/US/9603/education_summit/index.html.

Birkmire, K.J. (1993). Ninth grade proficiency, is it the best measure of learning: A survey of teacher attitudes. Retrieved from ERIC.

Brennen, R.L. (2001). Some problems, pitfalls and paradoxes in educational measurement. *Educational Measurement: Issues and Practices, 20*(4), 6-18.

Briod, M. (1989). A phenomenological approach to childhood development. In R. Valle & S. Halling (Eds.), *Existential-phenomenological perspectives in psychology: Exploring the breadth of human experience* (pp. 115-116). Norwell, MA: Kluwer Academic and Plenum Publishers.

Bruner, J. (1996). *The culture of education*. Cambridge, MA: Harvard University Press.

Child and Family Studies Department, University of Tennessee, Knoxville (2010). *Child and Family Studies core course requirements*. Retrieved from: http://cfs.utk.edu/ug/CoreCourses.html

Christie, K. (2003). One goal, different AYPs. Retrieved from: http://www.ecs.org/clearinghouse/44/68/4468.htm.

Christiansen, P., & James, A. (Eds.). (2000). *Research with children*. London: Falmer Press.

Cimbricz, S. (2002). State mandated testing and teachers' beliefs and practice. *Education Policy Analysis Archives, 10*, 2. Retrieved from: http://epaa.asu.edu/epaa/v10n2.html.

Coleman, J.S. (1972). The evaluation of equality of educational opportunity. In F. Musteller and D.P. Moynihan (Eds.), *On equality of educational opportunity* (pp. 140-161). New York: Vintage Books.

Corbett, D. and Wilson, B. (2001). *Listening to urban kids*. Albany: State University of New York Press.

Corbett, D. and Wilson, B. (1995). Make a difference with not for students: A plea to researchers and reformers. *Educational Researcher, 24* (5), 12-17.

Crisp, L., Williams, M. and Greenberg, K. (2006). *Digging deeper: The merits of phenomenological research in revealing the shared essence of teachers' perceptions of high stakes testing*. Presented at the American Educational Research Association Annual Meeting, San Francisco, California.

Darling-Hammond, L. (1989). Accountability for professional practice. *Teachers College Record, 91,* 59-80.

Darling-Hammond, L. (2007). The flat earth and education: How America's commitment to equity will determine our future. *Educational Researcher, 36*(6), 318-334.

Darling-Hammond, L. (1991). The implications of testing policy for quality and equality, *Phi Delta Kappan, 73*(3), 220-225.

Darling-Hammond, L. (2009). Teaching and the change wars: The professionalism hypothesis. In Hargreaves, A. and Fullan, M. (Eds.), *Change Wars.* Bloomington IN: Solution Tree.

Darling-Hammond, L. and A. Wise (1985). Beyond standardization: state standards and school improvement*. Elementary School Journal, 85*, 315-36.

Debard, R. & Kubow, P. K. (2000). Impact of proficiency testing: A Collaborative evaluation. Report prepared for Perrysburg (OH) Schools. Bowling Green State University and Partnerships for Community Action.

DeBray, E.H. (2006). *Politics, ideology, and education*. NY: Teachers College Press.

Department of Education (2003). Rules and regulations part II , 34 CFR part 200 Title I—Improving the academic achievement of the disadvantaged; final rule. *Federal Register, 68,* 236.

DeRosa, A. and Patalano, F. (1991). Effects of familiar proctor on fifth and sixth grade students' test anxiety. *Psychological Reports, 68*(1), 103-113.

Duncan, A. (2009). Higher standards, better tests, Race to the Top. Speech presented at the Governor's Education Symposium, Cary, NC.

Educational Management Group of Pearson (2010). Teachers' administration guidelines Retrieved from: http://www.state.tn.us/education/assessment/doc/ach_tam.pdf.

Ergene, T. (2003). Effective interventions on test anxiety reduction: A Meta-analysis. *School Psychology International, 24*(3), 313-328.

Evans, K.R., & Broemmel, A. (2008*). Listening to the teachers: Experiences in Success*

for All classrooms. Paper presented at the Annual Meeting of the National Reading Conference. Orlando, Florida.

Federal Education Policy and the States, 1945-2004: A Brief Synopsis (2006). *States' impact on federal education policy project*. Albany, NY: New York State Archives.

Ferguson, A. (2001). *Bad boys: Public schools in the making of black masculinity.* Michigan: The University of Michigan Press.

Figlio, D.N. (2006). Testing crime and punishment. *Journal of Public Economics, 90* 837-851.

Finn, C.E., Jr. (1990). The biggest reform of all. *Phi Delta Kappan, 71*(8), 584-592.

Fischer, C.T. (2006). Qualitative research methods for psychologists: Introduction through empirical studies. Amsterdam: Academic Press.

Firestone, W.A., Camilli, G., Yurecko, M., Monfils, L., & Mayrowetz, D. (2000). State standards, socio-fiscal context and opportunity to learn in New Jersey. *Education Policy Analysis Archives*, *8*(35).

Fleege, P. et al., (1992). Stress begins in kindergarten: A look at behavior during standardized testing. *Journal of Research in Childhood Education,7*, 120-26.

Freire, E. (1998). Teachers as cultural workers: Letters to those who dare teach. Boulder, CO: Westview Press.

Fuchs, D. and Fuchs, L.S. (1986). Test procedure bias: A meta-analysis of examiner familiarity effects. *Review of Educational Research, 56*(2), 243-262.

Fullan, M. (2009). Have theory, will travel: A theory of action for system change. In Hargreaves, A. and Fullan, M. (Eds.), *Change wars*. Bloomington, IN: Solution Tree.

Fullan, M. (with S. Stiegelbauer). (1991). The new meaning of educational change (2nd ed.) New York: Teachers College Press.

Fuller, F. (1969). Concerns of teachers: A developmental conceptualization. *American Educational Research Journal, 6*, 207-226.

Gadamer, H.G. (1966/1977). *Truth and method*. New York: Seabury.

Garcia, E.E. (1995). Educating Mexican American students: Past treatment and recent development in theory, research, policy, and practice. In J.A. Banks & C.A.M. Banks (Eds.), *Handbook for research on multicultural education*. New York: Macmillan.

Garder D.P., et. al (1983). *A nation at risk: The imperative for educational reform: A report to the nation and the Secretary of Education*. National Commission on Excellence in Education

Gitlin, A., Margonis, F. (1995). The Political aspect of reform: teacher resistance as good sense. *American Journal of Education*, 103, 377-405.

Gordon, S.P. and Reese, M. (1997). High stakes testing: Worth the price? *Journal of School Leadership, 7*, 343-367.

Graham, H. (1984). Surveying through stories, in Bell, C. & Roberts, H. (Eds.). *Social researching: Politics, problems, practice*, London: Routledge and Kegan Paul.

Greig, A. & Taylor, Jayne (1999). *Doing research with children*, London: Sage Publications Ltd.

Gunzenhauser, M.G., Hyde, A.M. (2007). What is the value of public school accountability? *Educational Theory, 57*(4), 489–507.

Hall, G.E., Georgi, A.A., & Rutherford, W.L. (1977). *Measuring stages of concern about the innovation: Manual for use of the SoC questionnaire*. Austin: Research and Development Center for Teacher Education, University of Texas.

Hall, G.S. (1903). Child study at Clark University: An impending new step. *American Journal of Psychology, XIV*, 96-106.

Hansen, V.K. (2006). *Globalization and the high school social studies curriculum: standards of learning in the southern states*. Paper presented at the Conference on Navigating the Globalization of the American South. UNC, Chapel Hill.

Hargreaves, A. (2009). The fourth way of change: Towards an age of inspiration and sustainability. In Hargreaves, A. and Fullan, M. (Eds.), *Change Wars*. Bloomington, IN: Solution Tree.

Hargreaves, A., Fullan, M. (Eds.) (2009). *Change Wars*. Bloomington, IN: Solution Tree.

Harris, C.R. (1991). Identifying and serving the gifted new immigrant. *Teaching Exceptional Children, 23*(4), 26-30.

Hatch, J.A. (1990). Young children as informants in classroom studies. *Early Childhood Research Quarterly, 5* (2), 251-264.

Heidegger, M. (1927/1962). *Being and time*. (Trans. J. MacQuarrie). New York: Harper & Row.

Heilig, J., & Darling-Hammond, L. (2008). Accountability Texas-style: The progress and learning of urban minority students in a high-stakes testing context. *Educational Evaluation and Policy Analysis, 30*(2), 75-110.

Hendrick, Henry (2000). The child as social actor in history, In Christiansen, P., & James, A. (Eds.) *Research with children: Perspectives and practices.* 36-61. London:

Falmer Press.

Hirsch, E.D. Jr. (1996). *The schools we need and why we don't have them*. New York: Doubleday.

Hoffman, J.V., Assaf, L.C., Paris, S.G. High stakes testing in reading: In Texas today, tomorrow? *The Reading Teacher, 54*(5), 482-492.

Hughes, S. and Bailey, J. (2002). What students think about high-stakes testing. *Educational Leadership, 59*, 74-76.

Igoa, C. (1995). *The inner world of the immigrant child*. New York: Martin's Press.

Jennings, J. (2003). Interview with Ben Wattenburg. *Think Tank*. Retrieved from http://www.pbs.org/thinktank/bio_1879.html.

Johnson, D.D. & Johnson, B. (2002). *High stakes: Children, testing, and failure in American schools*. Pennsylvania: Rowman and Littlefield, Inc.

Jones, B.D., & Egley, R.J. (2004). Voices from the frontlines: Teachers' perceptions of high-stakes testing. *Education Policy Analysis Archives, 12*(39).

Jones, G., Jones, B.D., Hardin, B, Chapman, L. Yarbrough, T., and Davis, M. (1999). The impact of high-stakes testing on teachers and students in North Carolina. *Phi Delta Kappan. 81*(3), 199-203.

Jones, G., Jones, B.D., Hargrove, T.Y. (2003). *The unintended consequences of high stakes testing*. Lanham, MD: Rowman and Littlefield.

Kessen, W. (1979). The American child and other cultural inventions*. American Psychologist, 34*(10), 815-820.

Klein, S.P., Hamilton, L.S., McCaffrey, D.F., & Stecher, B.M. (2000). What do test

scores in Texas tell us? *Education Policy Analysis Archives, 8*(49). Retrieved from: http://epaa.asu.edu/epaa/v8n49/.

Knox County Schools (2009). *State curriculum standards*. Retrieved from: http://knoxschools.org/modules/cms/pages.phtml?pageid=91051

Kogan, M. (1986). *Education accountability: An analytic overview*. London: Hutchinson.

Kohn, A. (2000). Burnt at the high stakes. *Journal of Teacher Education, 51*(4), 315-327.

Lattimore, R. (2005). African American students' perceptions of their preparation for a high stakes mathematics test. *Negro Education Review, 56*(2/3), 135-146.

Levine, B. (2009). *Reform without (much) rancor*. In Hargreaves, A. and Fullan, M. (Eds.), *Change Wars*. Bloomington, IN: Solution Tree.

Linn, R. (2000). Assessment and accountability. *Educational Researcher*, *29*, (2), 4-16.

Linn, R. (2005). *Test-based educational accountability in the era of NCLB* (CSE Report No. 651). Washington D.C.: National Center for Research on Evaluation, Standards, and Student Testing.

Lomax, R.G., West, .M.M., Harmon, M.C., Vintor, K.A., Madaus, G.F. (1995). *The Journal of Negro Education, 64*(2), 171-185.

Madaus, G.F., & Clarke, M. (2001). The adverse impact of high stakes testing on minority students: Evidence from 100 years of test data. In G. Orfield and Kornhaber, M.L. (Eds.). *Raising standards or raising barriers? Inequality and high-stakes testing in public education*. New York: The Century Foundation Press.

Manna, P. (2006). Control, persuasion, and educational accountability. *Educational Policy, 20*(3), 471-494.

Mantel, B. (2005). No Child Left Behind: Is the law improving student performance?

Congressional Quarterly Researcher, 15 (20), 469-492.

Mazzeo, C. (2001). Frameworks of state: Assessment policy in historical perspective. *Teachers College Record, 103*(3), 367-397.

MacNaughton, G., Smith, K., & Davis, K. (2007). Researching with children: The challenges and possibilities for building "child friendly" research. In Hatch, A. (Ed.) *Early childhood qualitative research.* New York: Routledge.

Macpherson, R.J.S. (1996). *Educative accountability: Theory, practice, policy and research in educational administration.* Tarrytown, NY: Elsevier Science Inc.

McClure, T. (2009). TVAAS 101. Retrieved from http://www.tennessee.gov/education/assessment/tools_resources.shtml.

McCracken, N.M. & McCracken, H.T. (2001). Teaching in the time of testing: What have we lost? *The English journal, 91*, 30-35.

McPhail, J. (1995). Phenomenology as philosophy and method: Applications to ways of doing special education. *Remedial and Special Education, 16*(3), 159-165.

Merleau-Ponty, M. (1941/62). *Phenomenology of perception*, trans. Colin Smith. New York: Humanities Press.

National Center for Education Statistics (2008)*The Nation's report card*. Retrieved from: http://nces.ed.gov/nationsreportcard/naep2008.asp.

National Education Association (2006). *School's who are risking sanctions increasing under NCLB.* Retrieved from: http://www.nea.org/index.html.

Nias, J. (1996). Thinking about feeling: The emotions of teaching. *Cambridge Journal of Education, 26*, 293-306.

Nichols, S. & Berliner, A. (2008). Testing the joy out of learning. *Educational Leadership,*

65(6), 14-18.

Nichols, S. & Berliner, A. (2005). *The inevitable corruption of indicators and educators through high-stakes testing*. (EPSL-0503-101).Tempe, AZ: Arizona State University, Education Policy Studies Laboratory.

Nieto, S. (1994). Lessons from students on creating a chance to dream. *Harvard Educational Review*, *64*(4), 392-426.

NYS Education Department (2006). Federal education policy and the states, Retrieved from: http://www.archives.nysed.gov/edpolicy/research/essay_Johnson_ESEA_distr.shtml

Noguera, P.A. (2007). How Listening to Students Can Help Schools to Improve. *Theory Into Practice, 46*(3), 205-211.

Nolan, S.N., Haladnya, T.M., & Haas, N. (1992). Uses and abuses of achievement test scores. *Educational Measurement: Issues and Practice, 11*(2), 9-15.

Ormrod, J.E. (2008) Educational Psychology: Developing learners, sixth edition. Upper Saddle River, NJ: Pearson.

Paris, S.G., Lawton, T.A., Turner, J.C., & Roth, J.L. (1991). A developmental perspective on standardized achievement testing. *Educational Researcher, 20*(5), 12-20.

Pedulla, J.J., et. al. (2003). Perceived effects of state-mandated testing programs on teaching and learning: findings from a national survey of teachers. Chestnut Hill,

MA: National Board on Educational Testing and Public Policy. Retrieved from ERIC database.

Pollio, H., Graves, R., & Arfken, M. (2005). Qualitative research methods. In F. Leong & J. Austin (Eds.), *The psychology research handbook: A guide for graduate students and research assistants*. Thousand Oaks: Sage.

Pollio, H.R., Henley, T.B., and Thompson, .J. (1997). The phenomenology of everyday life. New York: Cambridge University Press.

Polkinghorne, D.E. (1989). Phenomenological research methods, in R.S. Valle & S. Halling (Eds.), *Existential Phenomenological Perspectives in Psychology: Exploring the Breadth of Human Existence* (pp. 41-62). New York: Plenum Publishers.

Popham, W.J. (1981). *Modern educational measurement*. Englewood Cliffs, N.J.: Prentice-Hall, Inc.

Popham, W.J. (2001). The truth about testing: An educator's call to action. Alexandria, VA: ASCD.

Popham, W.J. (2007). Instructional insensitivity of tests: Accountability's dire drawback. *Phi Delta Kapan 89* (2), 146-153.

Popham, W.J. (2008). The Role of assessment in federal education programs. *Rethinking the Federal Role in Education*. Washington D.C.: Center on Education Policy.

Reisner, E., Atkin, M., Boruch, R., Linn, R., & Millman, J. (1982). *Assessment of the Title 1 Evaluation and Reporting System*. Washington D.C.: U.S. Department of Education, Office of Planning, Budget, and Evaluation.

Richardson, V. (1990). Significant and worthwhile change in teaching practice. *Educational Researcher, 19*(7), 10-18.

Roberts, H. (2000). *Listening to children: and hearing them*. London: Falmer Press.

Roderick, M. & Engel, M. (2001). The grasshopper and the ant: Motivational responses of low-achieving students to high stakes testing. *Educational Evaluation and Policy Analysis, 23,* (3), 197-227.

Rojcewicz, R. (1987). Merleau-Ponty and cognitive child psychology. *The Journal of Phenomenological Psychology*, 18, 201-222.

Salomon, G. (1984).Television is “easy” and print is “tough”: The differences investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology, 74*(4), 647-648.

Schunk, D.H., Hanson, A.R., Cox, P.D. (1987). Peer-model attributes and children’s achievement behaviors. Journal of Educational Psychology 79(1), 54-61.

Shepard (1993). *Setting performance standards for student achievement.* A report of the National Academy of Education Panel on the Evaluation of the NAEP Trial State Assessment: An evaluation of the 1992 achievement levels. Retrieved from ERIC.

Shirley, D. (2009). The music of democracy: Emerging strategies for a new era of post-standardization. In Hargreaves, F., Fullan, M. (Eds.), *Change Wars*. Bloomington, IN: Solution Tree.

Slavin, R. (2006). Evidence based reform and No Child Left Behind. *Teachers College Record* ID # 12889. Retrieved from: http://www.tcrecord.org.

Slavin, R. (2008). Perspectives on evidence based research in education – What works?

Issues in synthesizing educational program evaluations. *Educational Researcher, 37*(1), 5-14.

Smith, D.G. (1983) *The Meaning of children in the lives of adults: A Hermeneutic study*. Dissertation for Doctor of Philosophy, 1983, St. Thomas University, New Brunswick, Canada.

Spring, J. (2008). *The American school: From the Puritans to No Child Left Behind.* New York: McGraw Hill.

Stecher, B.M. (2002). Chapter 4: Consequences of large-scale, high stakes testing on school and classroom practice. In L.S. Hamilton, Stecher, B.M. & Klein, S.P. (Eds.). *Making sense of test-based accountability in education*. Santa Monica, CA: RAND.

Stiggins, R.J. (1994). *Student-centered classroom assessment*. New York: Merrill

Terzian, S. (2002). On probation and under pressure: How one fourth-grade class managed high-stakes testing. *Childhood Education, 78*(5) 282.

Tennessee Department of Education (2010). *Assessment, evaluation, and research division*. Retrieved from: http://www.tennessee.gov/education/assessment/index.shtml

Tennessee Department of Education (2010). *Test administration manual.* Retrieved from: http://www.tennessee.gov/education/assessment/index.shtml

Tennessee Department of Education (2007). *School profile*. Retrieved from: http://edu.reportcard.state.tn.us/pls/apex/f?p=200:1:2583943273772649::NO:::

Thomas, S.P., and Pollio, H.R. (2002). *Listening to patients*. New York: Springer Publishing Company.

Triplett, C.F., & Barksdale, M.A. (2003). Children's perceptions of high-stakes testing. *Journal of Research in Education, 13*(1), 15-21.

Triplett, C.F., & Barksdale, M.A. (2005). Third through sixth graders' perceptions of high-stakes testing. *Journal of Literacy Research, 37*(2), 237-254.

U.S. Department of Education (2002). *Testing for Results*. Retrieved from: http://www.ed.gov/nclb/landing.html.

U.S. Department of Education (2004). A guide to education and No Child Left Behind. Retrieved from: www2.ed.gov/nclb/overview/intro/**guide**/**guide**.pdf.

Valle, R.S.& Halling, S. (1989). *Existential-Phenomenological perspectives in psychology: Exploring the breadth of human experience*. New York: Plenum Press.

Valli, et.al. (2008). *Test driven: High-stakes accountability in elementary schools*. New York, NY: Teachers College Press.

van den Berg, R. (1974). Phenomenology and educational research. From Denton, D.E. (1974). *Existentialism and phenomenology in education.* New York: Teachers College Press.

van den Berg, R., & Ros, A. (1999). The permanent importance of the subjective reality of teachers during educational innovation: A concerns-based approach. *American Educational Research Journal, 36* (4)*,* 879-906.

Waber, D.P. and Bernstein, J.H. (1994). Repetitive graphomotor output in learning-disabled and nonlearning-disabled children: The repeated patterns test. *Developmental Neuropsychology, 10*(1), 51-65.

Wellhousen, K. & Martin, N. K. (1995). Preservice teachers and standardized test

administration: Their behavioral predictions regarding cheating. *Research in the Schools, 2*(2) 47-49.

Wilson, B.L., Corbett, H.D. (2001). *Listening to urban kids: School reform and the teachers they want*. Albany, NY: State University of NY Press.

Wheelock, A., Babell, D.J., Haney, W. (2000). What can students tell us about high-stakes testing in Massachusetts? Teacher College Record. Retrieved from: http://www.tcrecord.org. ID 10634.

Woolfolk, A. (2007). *Educational psychology*. Boston: Pearson Education.

Yakimoski-Srebnick, M.E. (2001). Security in a high stakes environment: The perceptions of test directors. Paper presented at the National Association of Test Directors' (NATD) Invited Symposium at the Annual Meeting of the National Council on Measurement in Education (NCME), Seattle, WA.

APPENDICES

Appendix A: Bracketing Interview

Bracketing Interview- October 8, 2007

I - This is a bracketing interview for Laura Crisp on October 8th, 2007. This is the second bracketing interview prior to collecting her data with children's experience with high stakes test. So say again your research question.

P – Tell me about the time you took the TCAP test.

I – So, what assumptions do you have about what you're going to hear?

P – Well, now that I've done the first preliminary interviews, um, I'm afraid I'm going to hear a cover story – what they think I want to hear. I actually – the first interviews I got – "they were really hard", or "I like them", or "my teacher really taught me well", or "I just want to do really well because I want to move on to the next grade". In all of them there was a bit of reticence to talk and just varying answers, even within interviews.

I - you said that judging from last time that you though they would tell you cover stories, what makes you think those were cover stories?

P – That's a good question. Because there were different opinions. And I'm thinking in particular of one little girl who said she really like them and then she said she didn't like them.

I – Were there other things you observed in the interviews that led you to believe they were cover stories?

P – Yes, I felt like they were telling me what I wanted to hear. I remember this one little boy, saying, leaning into the recorder, "And my teacher taught me really good." You know, like he was reading a script. "And I really want to go to college" So I could just tell, and maybe it wasn't a cover story so much as indoctrination. Just a result of being told what was right and what was wrong so I had kids

saying "staying quiet", and "taking their time" – you know things that their teachers would have said over and over again.

I – If they weren't telling cover stories what do you think they would be telling?

P – Um

I – or telling you the indoctrinated stuff

P – I think they – I was expecting the kids, because my group of focus was kids that didn't score well on the test. So I guess I had expected that they would say it was hard or that it was stressful. Particularly, it wasn't about taking the test but it was the ramification of the test. A couple of them said that they were afraid of going to the next grade, and they really wanted to do well on the test.

I – what evidence do you have that these children knew and or understood that they did not do well on the test?

P – Um, well, I don't have any evidence, really. I didn't ask them if they knew they did not do well on the test.

I – and one of the reasons I asked that is that you mentioned a moment ago, one of your assumptions was that because they didn't do well you expect them to you thought that they might say more that it was hard, that kind of hard. And I'm sitting here thinking, hm, what if they knew that they didn't do well?

I – on a related question, what would be the reason they would worry about going to the next grade based on this test.

P – That was something I wanted to look into whether that was something they made up or it was really tied to it.

I – were they all saying that?

P – two of them said it. I know in one case, the girl has been held back before, so I was thinking that was her motivation, but I wasn't sure about the other student.

I – Have you asked the school administrators how they determine when to retain a child?

P – I'm definitely going to make sure I find that out.

I - do you think teachers are encouraged to test students this will make a difference on whether you move on to the next grade.

P – I think this is something to explore as well. I couldn't imagine teachers lying to their students, and saying something like that if it's not true, especially, and I don't know when the teachers get the grades, but they aren't published until a week or two before school year starts, so I couldn't imagine kids beginning to find out they are not going to the next grade two weeks before school starts.

I – To my knowledge, they are almost always used to help make that decision. Because they are considered to be much more objective than teacher grading. And, so to make such a big decision, and this is an aside from the bracketing interview, but you need to know that it's often a way that they can justify the decision

P – I guess the only reason it might not be is because of the – they have so many other assessments besides TCAP and that they don't get the results back soon enough, from my knowledge – or maybe they just don't make them public til later but they have them.

I – In this district parents get them in September of the next year.

P – What were some other things that you learned about the last experience that you might apply to this experience?

I – um, I definitely learned a lot about the interviewing experience. When we were reading through [the transcriptions] with Pollio, it was just so enlightening. You know you really have to listen so carefully and just grab on to those words and ask back: "What do you mean by that, tell me more about that word." So I think I missed some opportunities that maybe I won't miss next time. Um, I

did a lot of, because they were children I asked Jessica on some recommendation on how to get more out of kids, and I used the recorder as an icebreaker, we played with the recorder a little bit. They loved hearing their voice on the recorder, which really helped them to be more comfortable with it as it can make them nervous. And then I also had a couple of toys that I brought that they could play with. And I also talked with them and got to know them on the way to the interview room. I wish I had known the kids personally. I think I would have gotten a lot more from them. And I think next time I'm going to spend more time explaining what I am doing. I touched on it a little bit but I want to make sure that I tell them that I am a student or part of the school. That I'm just a student at the University and I'm studying student's experience. And I think I used the word experience, and I'm not sure if they really understood that word so I am really going to spend some more time thinking about how I explain what I'm doing to ensure that we get more.

I – In most, it's been my experience that most teachers in schools make a big deal about TCAPS, have you had any understanding about how the test was introduced to them or how that may have affected the way they responded to you.

P – I've done a lot of reading up on it and I do know that some schools, especially schools like have rallies. And of course, they were told to sit quietly. And now that my son is at a school that didn't meet AYP the previous year, they are constantly doing TCAP prep, so you know they do tell them that they are really important. What I'd really like to do is spend some time on my dissertation exploring that so it is addressed in my study. It needs to be addressed.

I – do you think the students' families are saying anything to the students, peers or older peers?

P – that did come up in my last study. There were two twin boys in my study and they talked about their older brother and the older brother was the smart, talented one in the family, they were both in my study. And they were so sweet. And they were saying that they wanted to be like their brother.

They never, nobody, none of the kids mentioned their parents, and it probably is an assumption of mine that parents aren't talking to them a lot about TCAPS

I – did they talk about teachers?

P – they talked about teachers, one boy talked about how the teachers taught them real well in Math and that's why he did real well in Math. And he didn't do well in reading because he didn't know the words. But, when he did bad it wasn't the teachers fault, but when he did good, it was because of the teacher.

I – Laura, I would encourage you to really put a lot of thought into Michele's question. If the student's replies were protective of the teachers or as if they were feeling the need to ensure that people know, that may in fact be their experience with these tests.

Appendix B: Institutional Review Board (IRB) Approval Letter

THE UNIVERSITY of TENNESSEE
Institutional Review Board
Office of Research
1534 White Avenue
Knoxville, TN 37996-1529
Phone: (865) 974-3466
Fax: (865) 974-7400

03/29/2007

IRB#: 7280 B

TITLE: Listening to Students: The Lived Experience of Students Taking an Accountability Test

Crisp, Laura
Educational Psychology and Counseling
5717 Blossom Road
Knoxville, TN 37912

Greenberg, Katherine
Educational Psychology and Counseling
A 517 Claxton Complex
Campus - 3452

Your project listed above was reviewed and has been granted approval under Expedited review.

This approval is for a period ending one year from the date of this letter. Please make timely submission of renewal or prompt notification of project termination (see item #3 below).

Responsibilities of the investigator during the conduct of this project include the following:

1. To obtain prior approval from the Committee before instituting any changes in the project.
2. To retain signed consent forms from subjects for at least three years following completion of the project.
3. To submit a Form D to report changes in the project or to report termination at 12-month or less intervals.

The Committee wishes you every success in your research endeavor. This office will send you a renewal notice (Form R) prior to the anniversary or your approval date.

Sincerely,

Brenda Lawson
Compliances

Appendix C: Parental Informed Consent Letter

Parents Informed Consent Letter

Dear Parents:

Your child has been selected to participate in a research project. To complete this study, a graduate student from the University of Tennessee is conducting interviews with students. Knowing how your child did on the TCAP is also important to the study, so we will also be accessing your child's TCAP scores from last year (2006-2007 school year) and this year (2007-2008 school year). The research project is being conducted to better understand the experience of children who have taken the Tennessee Comprehensive Assessment Program (TCAP) test.

The interviews will be audio taped and will last as long as your child would like to talk about his or her experience taking the TCAP test but will last no longer than 45 minutes. Your child's TCAP scores and anything said during the interview will be kept in confidence. Neither the name nor any identifying information will be used in any reports although your child's words may be used to support the interpretation and analysis and appear in presentations and publications related to the research. At no time will the words be linked or traceable to your child.

I have read the above information. I have received a copy of this form.

☐ I agree to allow my child to participate in this study.

☑ I do not want my child to participate in this study

Participant's parent/guardian signature _____ Date _____

Child's Name _____

Investigator's signature _____ Date _____

Researcher's Contact Information:
Laura Crisp, Thornton Center
1801 Volunteer Boulevard
Knoxville, TN 37996
865-974-9876

Appendix D: Participant Assent Form

Appendix D

Assent Form (Participating Children)
Listening to Students: The Lived Experience of Students Taking an Accountability Test

Hello, (name of participant). Your parent or guardian said that it would be o.k. for you to help me by answering some questions. All that you have to do is answer a question about what it was like for you when you took the TCAP test. Then, I may ask additional questions such as, "Can you say more about 'that'?" I think that what we will learn from you will help me and others understand your experience better. Are you willing to help with this project? (Child's response). Great! I think you will find that you really will really enjoy sharing your experiences. The interview won't last longer than 45 minutes but, if you decide that you don't want to talk anymore, all you have to do is tell me. You can just say, "I don't want to talk about this anymore." Okay? (Child's response). I will be turning this recorder on so that we can record what we talk about. Okay? (Child's response).

I really appreciate your help! Are you ready? (Child's response). Let's begin.

APPROVED
By Brenda Lawson
MAR 2 9 2007

Appendix E: Research Team Pledge of Confidentiality

Appendix C

Research Team Member's Pledge of Confidentiality

As a member of this project's research team, I understand that I will be reading transcriptions of confidential interviews. The information in these transcripts has been revealed by research participants who participated in this project on good faith that their interviews would remain strictly confidential. I understand that I have a responsibility to honor this confidentially agreement. I hereby agree not to share any information in these transcriptions with anyone except the primary researcher of this project, his/her doctoral chair, or other members of this research team. Any violation of this agreement would constitute a serious breach of ethical standards, and I pledge not to do so.

Research Team Member	Date
	11/19/07
	11/19/07
	11/19/07
	11/19/07
	11/19/07
	11/19/07

Appendix F: Potential Themes Presented to Research Committee

Different

Int 1: Lines 12-13- it was kind of weird because I had never taken one before, never taken such a long test

Int 1: 47-49 - So I would have to look it up in my book, but in the TCAP, I can't do that. Like I knew what the word meant, I just didn't know that that was the word for what it meant. That kind of confused me

Int 1: 122-124 - it's kind of *weird* everytime, because you have to be so quiet and stuff, if you kind of get loud, it's kind of bad. But, it's kind of good when you take a break, because its um, you have to be quiet and stuff

Int 1: 198-199 - but you can't do that at the test, look at your friend, because they think that you are cheating.

Line 88-89 **-** well what was good about it, they told us when we had five, ten, fifteen minutes of it, so that was good about it.

Int 2: Lines 27-28 - I'm making it a really big deal in my head,

Int 4: Line 6 - I studied for the TCAP to get a good grade on it.

Int 4: Line 32 - its just peace and quiet so

Int 5: Line 11 - it just felt like, another test, but like, more important

Int 5: Line 51 - well hard is like, I mean so many pages and stuff…

Int 9: Line 120-124 - they could only read the question for us cause like if they um, if they told us the answer it would be cheating and that's why we had another teacher in our room too. And, um, because most of the classes they switched teacher because um, because they think that like they'll cheat and give you the answer. Our teacher told us that she said we was lucky because a lot of teachers, they switch, but we just had another teacher.

Stress

Int 2: Lines 41 - it was also fun

Int 2: Line 45 - **well the social studies, it's fun, because I like social studies, answering questions about social studies. That's my favorite subject so it's always fun for me**

Int 2: Lines 53-54 - well, like the questions, some of them were fun to work out but some of them were hard to work out

Int 2: Lines 6-7 - – It was, kind of hard, cause you're stressed out because you know this stuff so you're making it hard,

Int 2: Line 63 - but I was makin it like so hard on myself
Int 2:Line 11-13 - it wasn't like sleepy tiring it was like, you just, you're so stressed out about it because there's nothing really to be stressed out about you just get so tired from being stressed out you just want to lay down and go to sleep.

Int 2: Line 70-71 - , so then I just made myself think it was hard so most of stuff, so then it kind of was hard

Int 2: Line 54 - when you try hard, you like, its like you're trying hard to hold your breath

Int 4: Line 38 - well, sometimes I'm nervous, I don't want to miss that many, so

Int 4: Line 78-79 - Well, sometimes, like when I get nervous, I'll forget some things. I try to think back about it and how the teacher taught it to us.

Int 7: Line 80-82 - It wasn't like, it was like you're falling asleep and you're trying hard and stuff and you pick up the pencil just get really nervous and when you are nervous it's like um, you wonder what your surprise is going to be for Christmas

Int 7: Line 95-99 - when you're nervous I just feel like um, I just feel like you are at a football game and you want to win really really bad and it's like when you're nervous it's like your nervous when the roller coaster is really really high and you're nervous because it's really high and you are afraid of heights and when you are going down a roller coasters it kind of drops your stomach and that's what it feel like

Int 7: Line 108-110 - when you are half way through the test it's when you start to get used to it and you aren't nervous anymore and um, when you are not nervous it kind of feels good, cause it's like you are the smartest kid in the world.

Int 8: Line 11-12 - , that was my first year last year and I didn't know what it would be like to take it

Int 9: Line 83-84 - , is this going to be hard or is this going to be easy? Because I never took it before

Degree of Difficulty

Hard and Easy

Int 1: Lines 15-17 - Some of the questions were kind of hard and some of them were kind of easy. It would depend on whether or not I knew it really well or didn't know it very well. So, if I knew it very well I would go really fast

Int 2: Lines 6-7 - – It was, kind of hard, cause you're stressed out because you know this stuff so you're making it hard,

Int 2: Line 63 - but I was makin it like so hard on myself

Int 2: Line 70-71 - , so then I just made myself think it was hard so most of stuff, so then it kind of was hard

Int 3: Line 5 - it wasn't that hard

Int 3: Line 9 - it was a little bit easy

Int 3: Line 13-14 - I missed one question, I think, of the, about that expanded form, how you got to write that number thing, expanded form, and that's all

Int 3: Line 40 - the grammar, I missed three…

Int 5: Line 7 - I didn't really have trouble with it

Int 5: Line 45-6 - I mean it was kind of hard, but then again it wasn't

Int 5: Line 51 - well hard is like, I mean so many pages and stuff…

Int 6: Line 6 - I kind of struggled a little bit,

Int 6: Line 14 - there aint much to it

Int 6: Line 67 - um, it was kind of difficult and some of it was easy for me.

Int 7: Line 29-30 - yeah, you have to really try hard and um try hard to figure out the problem. If you don't know it, the problem you can go ahead and come back to it

Int 7: Line 49-50 - when we done it it was really hard, I still tried my hardest, give 110% and you always have to try hard and keep trying and trying til you get it.

Int 7: Line 65 - it was kind of like medium and hard, it wasn't that easy

Int 7: Line 69 - sometimes I had to just guess 'cause it was just too hard.

Int 8: Line 5-6 some questions were kind of easy some I had to….some were kind of easy and it wasn't *too* hard

Int 8: Line 6-7 - when I first took it, I thought it would be kind of hard because, but….

Int 8: Line 16 - it was like o.k. for some of it

Int 8: Line 29-31 - , the science questions, they like, they were kind of easy I knew what they meant. And social studies was o.k., but some things I didn't really get on the social studies

Int 8: Line 45 - it was like, it wasn't really challenging because it was already what we learned

Int 9: Line 5 - It was kind of hard, but it was kind of easy

Confusing

Int 1: Lines 22 - some of the problems I couldn't understand

Int 1: Lines 39-41 - like I knew what it was, but I just didn't know that was the word for it. So I was kind of confused like that.

Int 1: 47-49 - So I would have to look it up in my book, but in the TCAP, I can't do that. Like I knew what the word meant, I just didn't know that that was the word for what it meant. That kind of confused me

Int 5: Line 72 - it just like we read it, like a problem or something, it just seemed confusing
Int 5: Line 90-91 - it just didn't make that much sense to me, it didn't go through my mind right or something.

Int 8: Line 29-31 - , the science questions, they like, they were kind of easy I knew what they meant. And social studies was o.k., but some things I didn't really get on the social studies

Int 8: Line 41 - I didn't understand the question.

Fun and boring

Int 2: Lines 41 - it was also fun

Int 2: Line 45 - well the social studies, it's fun, because I like social studies, answering questions about social studies. That's my favorite subject so it's always fun for me

Int 2: Line 95-96 - it kind of makes you kind of tired because you get bored and you're kind of sleepy and stuff from doing all those problems and your hand starts hurting and…

Int 2: Lines 53-54 - well, like the questions, some of them were fun to work out but some of them were hard to work out

Int 6: Line 17 - it was boring

Int 6: Line 20 - and it was no fun, I mean it was fun, but sometimes it got a little boring

Int 7: Line 5 - It was fun and it was really good and all you have to really do is concentrate

Int 7: Line 9-12 **- It was fun because in Math, well I like doing math, subtracting and adding, and social studies, it was fun to learn about new stuff. And in science, I like to learn about**

plants and I like, um, to do projects and writing, I like to write about books and I like to read books

Int 7: Line 75-76 - when we took the TCAPs I was excited because we got the highest test scores and we had like an ice cream party and it was fun.

Int 9: Line 19 - it was kind of fun

Likes and Dislikes

Int 2: Line 45 - **well the social studies, it's fun, because I like social studies, answering questions about social studies. That's my favorite subject so it's always fun for me**

Int 2: **Line 77-78 - when I don't really like stuff, I don't really remember it that well**

Int 2: Line 88-89 **-** well what was good about it, they told us when we had five, ten, fifteen minutes of it, so that was good about it.

Int 3: **Line 18 - I don't like math**

Int 6: Line 47–50 - **Well the reading part is where I get to read stories and I love to read stories and answer the questions and circle which one is the right answer. And the social studies questions you really entertain you 'cause it tells you what the people did and how they did or what they did in the Civil War.**

Int 7: Line 9-12 **- It was fun because in Math, well I like doing math, subtracting and adding, and social studies, it was fun to learn about new stuff. And in science, I like to learn about plants and I like, um, to do projects and writing, I like to write about books and I like to read books**

Int 8: Line 20 – 21 - **well I think the easiest was math and the reading because those are my favorite subjects**

Int 2: Line 89 - And snack was good too.

Int 2:Line 94-96 - So the snack was good. That was a good thing we got to look forward when we finished the test and we turned in the packet we got a snack

Int 5: Line 105 - We got snacks, and before the test we got like peppermints. And so I like that.

Int 7:Line 38-40 - the teacher allowed us to have snack and um, we had snack, we had juices and stuff and that helped us to concentrate more and sometimes on the TCAP we got to take peppermints and that also helps us concentrate

Int 9: Line 13-14 - we had snack, and we had, we got to drunk our water and we got some juice and stuff

Self-Awareness/Motivation

Int 1: Lines 15-17 - Some of the questions were kind of hard and some of them were kind of easy. It would depend on whether or not I knew it really well or didn't know it very well. So, if I knew it very well I would go really fast

Int 2: Line 45 - **well the social studies, it's fun, because I like social studies, answering questions about social studies. That's my favorite subject so it's always fun for me**

Int 2: **Line 77-78 - when I don't really like stuff, I don't really remember it that well**

Int 3: **Line 18 - I don't like math**

Int 4: Line 83-84 - , I just, like, take a minute to think about it, then I'll remember it. I try to have very much self-confidence as I can.

Int 6: Line 47–50 - **Well the reading part is where I get to read stories and I love to read stories and answer the questions and circle which one is the right answer. And the social studies questions you really entertain you 'cause it tells you what the people did and how they did or what they did in the Civil War.**

Int 7: Line 9-12 **- It was fun because in Math, well I like doing math, subtracting and adding, and social studies, it was fun to learn about new stuff. And in science, I like to learn about plants and I like, um, to do projects and writing, I like to write about books and I like to read books**

Strategies

Int 4: Line 58-60 - social studies, some things we try to go over, but we don't even get to that much, so I just try to make the right guess and stuff. And usually we do more Math than anything, and reading. Mostly reading

Int 4: Line 78-79 - Well, sometimes, like when I get nervous, I'll forget some things. I try to think back about it and how the teacher taught it to us.

Int 4: Line 83-84 - , I just, like, take a minute to think about it, then I'll remember it. I try to have very much self-confidence as I can.

Int 6: Line 6-7 - I skipped it and went on to another one and I then checked over in case I missed one and made a mistake

Int 6: Line 38–40 - some reading questions will be in there that you have on your work and if you study for that work you remember to go for that question on the TCAP test.

Int 7: Line 5 - It was fun and it was really good and all you have to really do is concentrate

Int 7: Line 29-30 - yeah, you have to really try hard and um try hard to figure out the problem. If you don't know it, the problem you can go ahead and come back to it

Int 7: Line 34 - and always give 110%...

Int 7: Line 49-50 - when we done it it was really hard, I still tried my hardest, give 110% and you always have to try hard and keep trying and trying til you get it.

Int 7: Line 69 - sometimes I had to just guess 'cause it was just too hard.

Int 9: Line 64 - some questions I didn't know, like I kind of just guessed.

Int 9: Line 68-72– because I didn't know the answer, and the teacher said last year, like, if you don't know the answer, just like, pick C, because that's what they should put the answer on. And sometimes I just guessed and sometimes I just put C. And the easy ones were the ones I knew good and the ones I could do good and um some, like on the reading, some I got if I go look at the story.

Vita

Laura Rutherford Crisp received a Bachelor of Arts degree in Political Science at the College of Charleston in 1999. She was a journalist for the School's Plus section at the *Post and Courier*, the local daily newspaper in Charleston, writing positive pieces about the schools in the area. After having her son, Ethan, she moved back to her hometown, Knoxville, TN, and continued writing in the business sector, educating clients about operating room software and coordinating conferences and tradeshows at GE Medical Systems Information Technologies in the Marketing Department.

Laura decided to go back to school to pursue her passion and enrolled in the Ph.D. program in Applied Educational Psychology. Working as a graduate assistant at the Thornton Center, a student-athlete support center, she coordinated academic support for students in the Special Needs Program (SNP). She also began to directly tutor college students with special needs as an SNP mentor. With her colleague, Michele Williams, she co-developed and taught a study hall that introduced Cognitive Education Approach (CEA) to student-athletes struggling academically. She also taught CE 212, a summer course for upcoming freshman athletes and taught a research studies course for the McNair program.

Laura continued adding to her 'tool belt' of cognitive and learning therapies and began working as a Learning Specialist with the Psychoeducational Network in 2005, where she continues to work. During this time, she also worked with Webb School of Knoxville as their Learning Specialist and Learning Center Director. Her next career goal is to seek out and acquire effective, innovative learning therapies to improve cognitive

performance for those with learning differences and psychological challenges, and bring this expertise to the Knoxville area.